MASTER OF HUMAN INTERFACE TECHNOLOGY

GESTURE BASED INTERACTIONS FOR AUGMENTED VIRTUAL MIRRORS

UC
UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

JONATHAN WONG
Human Interface Technology Laboratory New Zealand
College of Engineering
University of Canterbury
2015

Supervisory Team: Dr. Christoph Bartneck (Primary), Dr. Gun Lee and Professor Mark Billinghurst.

HITLabNZ
Human Interface Technology Laboratory New Zealand
www.hitlabnz.org

This thesis is the work of JONATHAN WONG, in fulfilment of the Master of Human Interface Technology degree at the Human Interface Technology Laboratory New Zealand (HITLab NZ), under the College of Engineering at the University of Canterbury.

ABSTRACT

In this research we explore the development of a gesture interface for an interactive virtual mirror display. We follow a user-centred approach and utilised interaction design methods. Through the course of designing and developing the interface, in order to investigate natural gesture interaction methods we conducted three user studies : a guessability study, a mapping study, and a target study.

Twenty participants were recruited in a guessability study to discover user-defined gestures for common interaction tasks. For pointing tasks, 80% of users used a hovering gesture, with the rest using directing gestures. Swiping gestures were used in the majority (81%) of scrolling tasks. For selection tasks, a mixture of waiting, tapping and grabbing gestures were observed. When choosing between dichotomous options, a mix of thumbs up/down and directing gesture were used. Results suggests that for some tasks, designing a system to support multiple interaction methods may be beneficial for usability.

Next, we recruited ten participants for a mapping study evaluating the use of a non-linear mapping of the interface cursor. Our interface showed potential as an interaction method for where a direct mapping method was unsuitable, such as when a target was out of reach.

Finally, twenty participants took part in a target study where we evaluated a hovering, extension and pointing interface. The hovering interface was ranked as the easiest and most fun to use, followed by the extension interface, although task completion time between the hovering and extension interfaces showed no significant differences. The pointing interface had the longest task completion time and was ranked the lowest in ease of use and overall preference.

In conclusion, the use of a user centred approach resulted in the development of three gesture interfaces utilising a hovering, extension and pointing interaction method that were felt to be natural and intuitive.

ACKNOWLEDGEMENTS

I have learned an immeasurable amount during the course of my research and wish to acknowledge my supervisors- Professor Mark Billinghurst, Dr. Gun Lee and Dr. Christoph Bartneck for their guidance and advice. Special thanks also to my friends, colleagues and faculty members at the HITLab NZ and College of Engineering for making each day an enjoyable experience.

This work was partly supported by the ICT R&D program of MSIP/IITP. [No.15501-14-1016, Instant 3D object based Join&Joy content technology supporting simultaneous participation of users in remote places and enabling realistic experience]

CONTENTS

1 INTRODUCTION

The use of gestures is becoming widely adopted as a mode of controlling a computer system, with the promise of a human-machine interaction where users can directly, and effortlessly, interact with a system. This thesis explores the use of gestures for interacting with large screen displays and the development of such gestural interfaces.

1.1 GESTURE INTERFACES

The use of depth sensing technology is seeing increasing popularity, with the introduction of consumer products like the Leap Motion [Leap Motion, 2015], PlayStation Eye [Sony, 2013] and the Microsoft Kinect [Microsoft, 2015d]. These devices have the ability to track a user's body movement and consumer understanding is maturing to a point where users are starting to realise their potential use beyond its novelty and gaming applications.

A risk to the growing number of gesture interfaces however is the lack of any interaction standards, leading to a fragmented user experience [D. Norman and J. Nielsen, 2010]. Depending on the situation, a circular motion of your hand might change the channel on your TV, increase the volume of your car's radio or even adjust the brightness on your laptop screen. The lack of a consistent user experience between gestural interfaces presents a usability issue that could be detrimental to its adoption. The motivation for manufacturers to develop gestures that are patentable, or for researchers to develop gestures than can be easily recognised by a computer, can be at the expense of the user experience, threatening the core principles of the use of gestures as a natural user interface.

This research aims to discover the natural gestures used when interacting with a large screen display. While there has been previous research into the area, these past work focused on gestures developed or proposed by the researchers. By following a user-centred design approach, we hope to highlight the importance of developing a gesture interface from the user's perspective.

1.2 INSPIRATION

Inspiration for this research came with the release of the movie, Minority Report [Spielberg et al., 2002]. In a scene from the movie, the protagonist is seen wearing illuminated tracking gloves (Figure 1) which allowed him to interact with a large interactive display; using swiping and pinching gestures to manipulate objects. The effortless interaction displayed in that scene was a stark contrast to the clunky computer keyboard and trackball mouse commonly in use at the time, and provided a glimpse into the future of human-computer interaction.

Figure 1: A scene from the movie Minority Report (2002) where the protagonist interacts with a gesture interface using tracking gloves. Picture from [Caron, 2015].

1.3 THE PROCESS

In order to explore the use of gestures for interacting with large screen displays we followed an iterative design process (Figure 2). The process begins with a literature review, where we summarise the previous work done in the field. We can then design an experimental study and develop a prototype to test our ideas. Following from this we enter an evaluation phase to identify potential issues in our design and gather data on its performance. We can undertake several design iterations, modifying our approach until the performance is ideal.

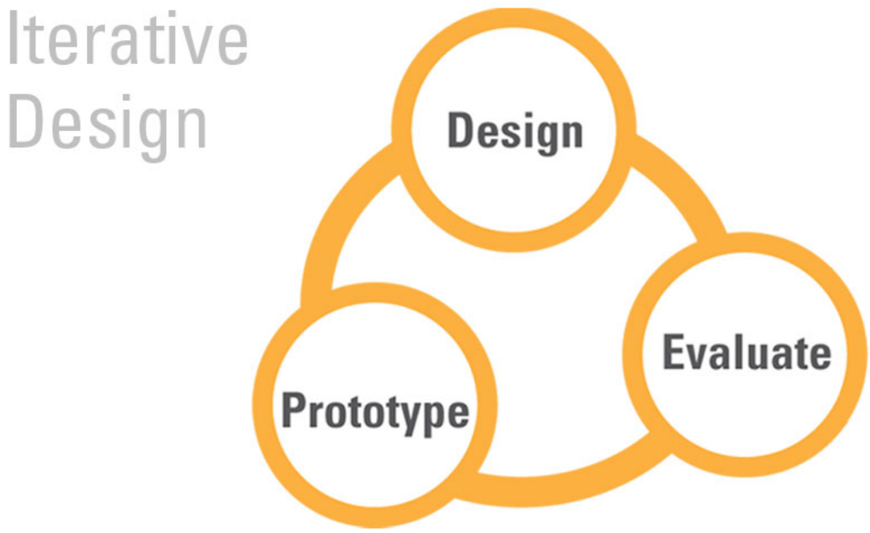


Figure 2: An overview of our research process, based on an iterative design model from [Human-Computer Interaction Institute, 2011].

In the next section, Chapter 2, we summarise the background research and the existing commercial work done in the field of gesture interfaces.

Following that we provide an overview of the design considerations that should be considered when developing a gesture interface in Chapter 3 and review the hardware/software tools that we used for the implementation of our prototype(s) in Chapter 4. In Chapter 5 we describe our initial experiment, a guessability study, to discover user-defined gestures for interacting with interactive public displays. The results are summarised in Chapter 6. In Chapter 7 we reiterate our design to explore various cursor mapping methods in a mapping study, before reporting the results in Chapter 8. We undergo another iteration in a target study (Chapter 9), evaluating three prototype interfaces. The results from our target study is summarised in Chapter 10, before concluding our research in Chapter 11, summarising our findings and potential areas for future work .

2 BACKGROUND

With the growing ubiquity of mobile devices came the popular use of touchscreen interfaces that used fingers for input, in place of the traditional mouse and keyboard. Initial touchscreen interfaces ported the traditional desktop WIMP (Windows, Icons, Menus and Pointers) model directly onto touchscreen devices. As a result, these early interfaces provided a poor user experience. However, devices like the iPhone and iPad took a different approach. Following a user-centred approach, their interface was distinct from the tradition WIMP layout, specifically with:

- the removal of the mouse cursor
- replacing desktop icons with larger, touch-friendly icons
- using full screen applications, not minimised into a window frame.

These changes provided users with a richer and more 'native experience' as they adapted to a new mode of interaction. As we move towards a new mode of human-computer interaction, the use of gestural interfaces, the lessons learned from our transition to touch interfaces- i.e. the benefits of a user centred approach should be considered .

Figure 3: Early tablets directly ported the WIMP interfaces (left) onto touchscreen devices, while the iPad (RIGHT) utilised an interface designed specifically for touch interactions. Left picture from [Microsoft, 2015a], right picture from [HotHardware, 2015].

There is a progression in human-computer interaction towards the use of natural user interfaces (NUI). These systems incorporate multi-modal interactivity, e.g. allowing speech and gesture inputs and promising a direct interaction with computers, free from intermediate hardware or controllers [Lee et al., 2012]. Such interfaces are commonly used in the form of interactive large screen displays, and thus we focus our research into the gesture interactions with these systems.

2.1 WHAT IS A GESTURE?

A gesture, for the purpose of designing gestural interfaces, can be defined as any deliberate movement of the body for the intentional communication of an idea or action [Rico and Brewster, 2010]. With this definition we can exclude subconscious or behavioural gestures- like scratching an itch or readjusting your reading glasses.

There are various gesture classifications. For example, [Cassell, 1998] proposed a classification for the human-computer interface community, adapted from [McNeill, 1992], who categorised gestures into five groups: beat gestures, deictic gestures, iconic gestures, metaphoric gestures and cohesive gestures. A brief summary of their features is as follows:

- Beat gestures are rhythmic, repetitive movement of the hands, e.g. flicks.
- Deictic gestures are pointing gestures. The pointing may be targeted at a physical reference, e.g. pointing at an object, or may be abstract, e.g. pointing at a point in space (see Figure 4).
- Iconic gestures are associated with expression, e.g. holding your hands apart when describing something big, or flapping your hands by your side when describing a bird's flight.
- Metaphoric gestures represent abstract concepts with no physical form. They are not necessarily related to a metaphor, e.g. using a rolling motion with your hands when describing a meeting that "went on and on."
- Cohesive gestures are those that are related by concept but temporally separated. For example, one might start gesturing a path in the air but get interrupted mid-way. Their hands may relax by their side during the interruption, but subsequently return to the last position to complete the gesture.

Figure 4: (Left) A deictic gesture using the index finger, (Centre) A deictic gesture using an open palm, (Right) An iconic gesture. Picture from [Yoshioka, 2005].

The use of gestures for communication predates the invention of the computer. As such, there is a rich history of gestures for use in different industries and can vary according to context. While gestures may be classified, some gestures can consist of more than one type, and classification systems are often subject to change and agreement. Still, they have proven useful when developing a gesture vocabulary. Examples of gesture vocabularies for in various situations are shown in (Figure 5) and (Figure 6), showing the possible gesture sets used when scuba diving, or in tactical situations, respectively.

2.2 TECHNOLOGICAL APPROACH TO GESTURES

Early development of gesture interfaces approached the problem from a technical perspective, defining gestures that were easily recognised by a computer [M. Nielsen et al., 2003]. While this technological approach may be easily implemented, usability testing have found that the gestures were often illogical, physically stressful, and sometime even physically impossible to perform for some users (see Figure 7). Another disadvantage of such an approach is that it tends to preserve the status quo, restricting interactions to the limits of current technology [M. Nielsen et al., 2003].

2.3 USER CENTRED APPROACH TO GESTURES

[Rico and Brewster, 2010] found gestures to be most comfortable if they were subtle, the movement was enjoyable, and if the gestures were similar to existing technology or everyday actions. Uncomfortable gestures were found to look weird, were physically uncomfortable, and/or disruptive to normal behaviour. Some of the best gestures are those that matches a user's natural behaviour, complementing an action the user is already performing. An example given by [Saffer, 2008] is the introduction of radio-frequency identification (RFID) cards for use as a method of payment in public transportation. While the original intention was that the RFID cards be taken out and scanned, it was noticed that users would often just leave their RFID cards in their bag and make a swinging motion of their bag across the card reader as they pass, adapting the interaction into a casual, natural gesture.

2.4 GESTURE USABILITY

Previous work with gesture interfaces have focused on the learnability of gestures pre-defined by the researcher, rather than a gesture defined by the user. [Ackad, Kay, et al., 2014] developed an interface using gestures inspired by sign language (see Figure 8), but quickly found that participants had trouble learning them and often needed a demonstration from the researchers, even though their interface also had a tutorial.

[Hespanhol et al., 2012] observed users performing a selection task using five different gesture types: pushing, dwelling, lassoing, grabbing and enclosing. For the pushing gesture, users had to simulate pushing a physical button in midair. The dwelling gesture consisted of hovering their hand over a selection. The lassoing gesture required users to circle around their selection. The grabbing gesture simulates the grasping of an object, while the enclosing gesture closely follows the grabbing one but used two hands to grab the selection. Their results found the dwelling gesture was the most intuitive. The pushing gestures, commonly used in touch-based interfaces,

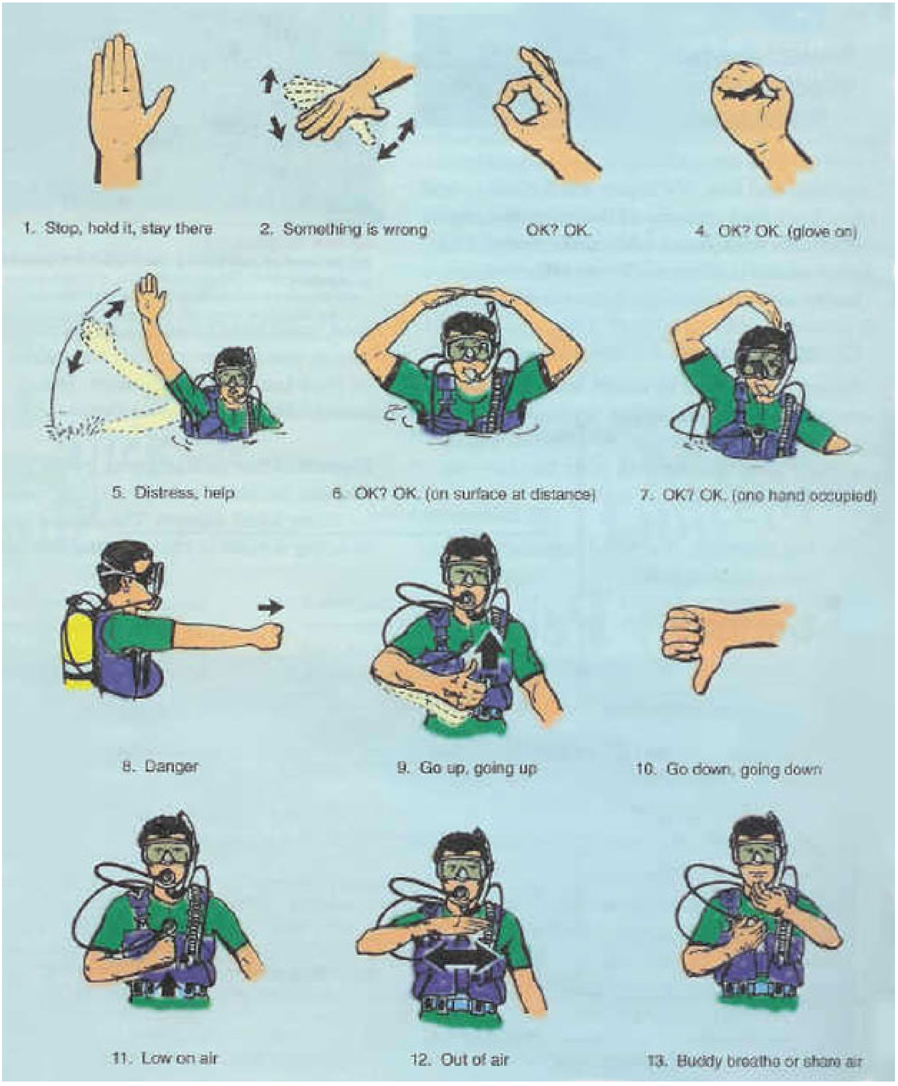


Figure 5: Gestures used in scuba diving. Picture from [Scuba, 2015].

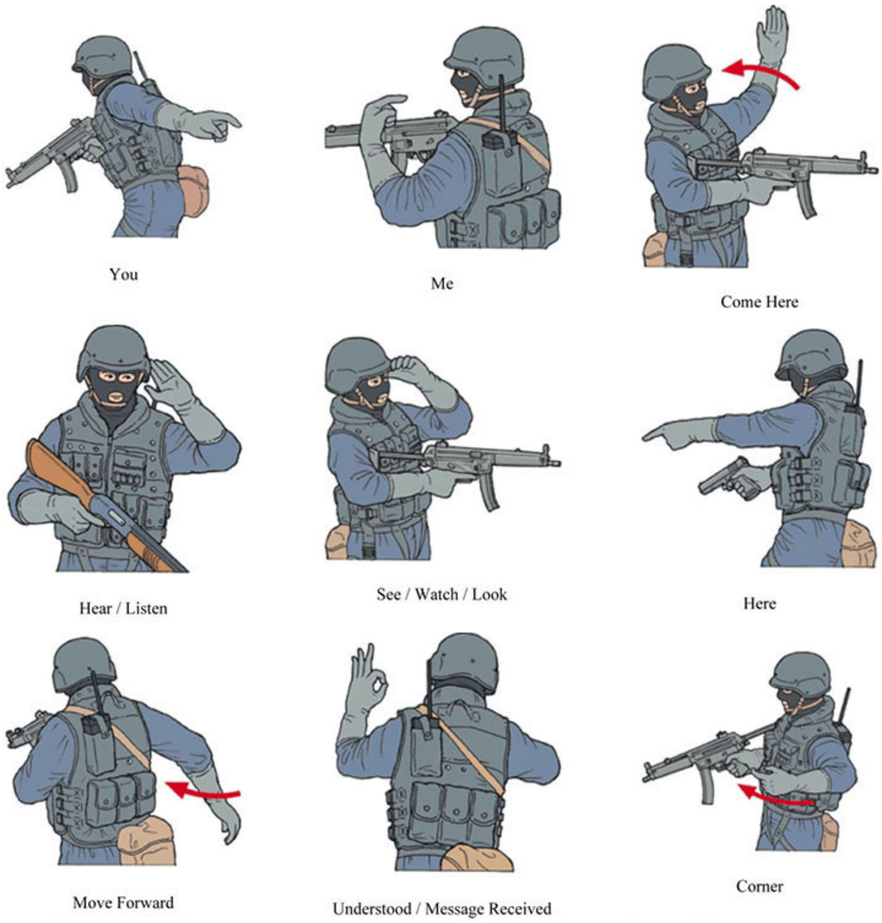


Figure 6: Gestures for tactical situations. Picture from [Uscrow, 2015].

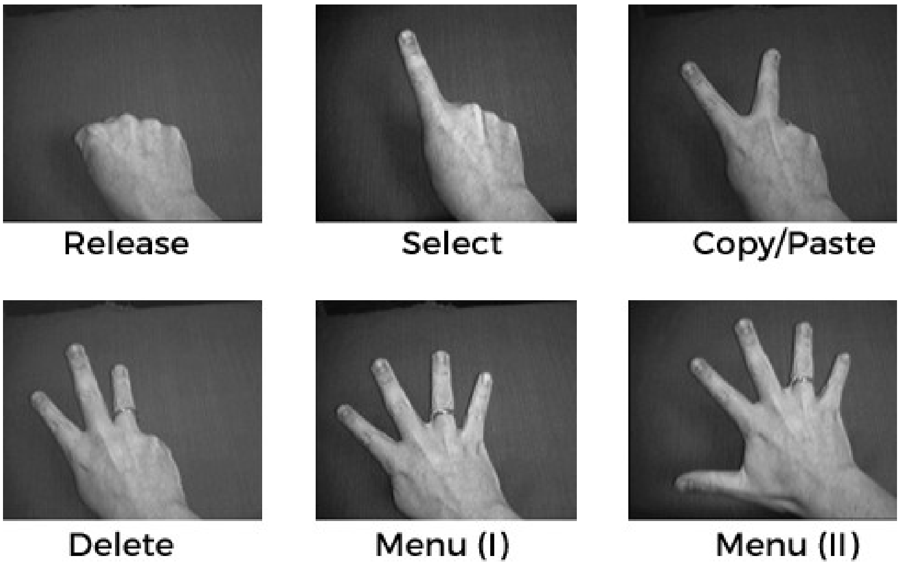


Figure 7: Technically developed gestures and their intended actions. Picture adapted from [M. Nielsen et al., 2003]

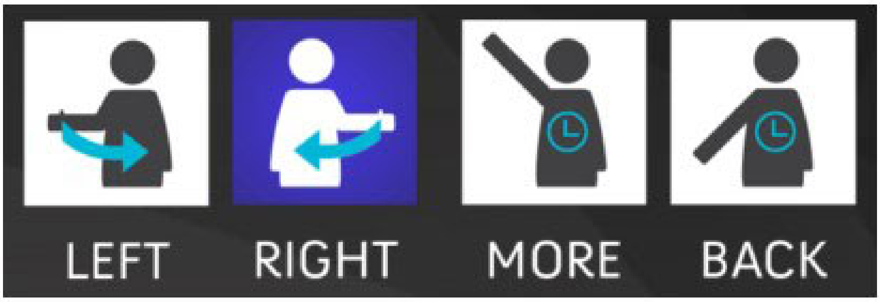


Figure 8: A gesture interface vocabulary inspired by sign language. Picture from [Ackad, Kay, et al., 2014].

was not favoured by participants, presumably due to the lack of tactile feedback. (Figure 9) summarises the gestures used in their study.

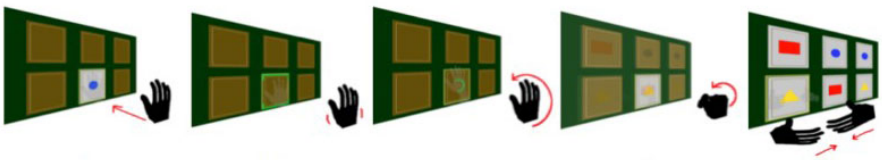


Figure 9: Gestures for selecting (a) pushing (b) dwelling (c) lassoing (d) grabbing (e) enclosing. Picture from [Hespanhol et al., 2012].

[Nancel et al., 2011] examined the navigation tasks of panning and zooming on a gesture interface. They found linear gestures not only faster, but also more preferred by participants than circular gestures (see Figure 10).

[Jakobsen et al., 2013] studied the use of proxemics (the user's body position) to interact with a large screen display. For examples, a user could move closer to a screen to enlarge the content, or move left/right to navigate a map on the screen (see Figure 11). Content may also change depending on the direction a user is facing relative the screen. The use of proxemic

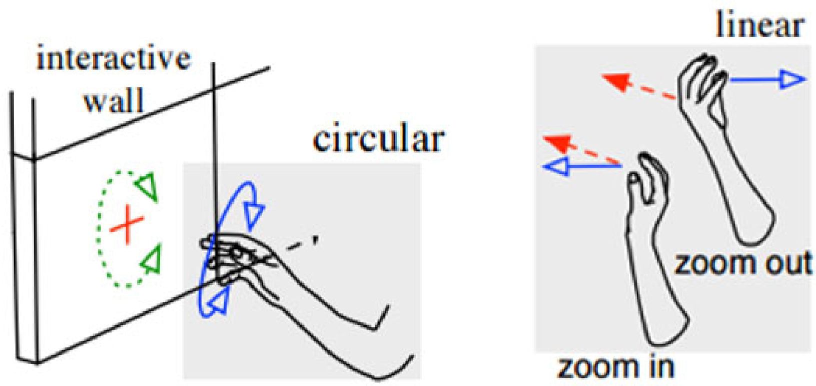


Figure 10: Circular gesture (left) vs linear gestures (right) for panning and zooming tasks. Picture from [Nancel et al., 2011].

interactions was found to be useful and natural, although some users were unsure about the control threshold and wanted to freeze the content from continuously responding to their body movement.

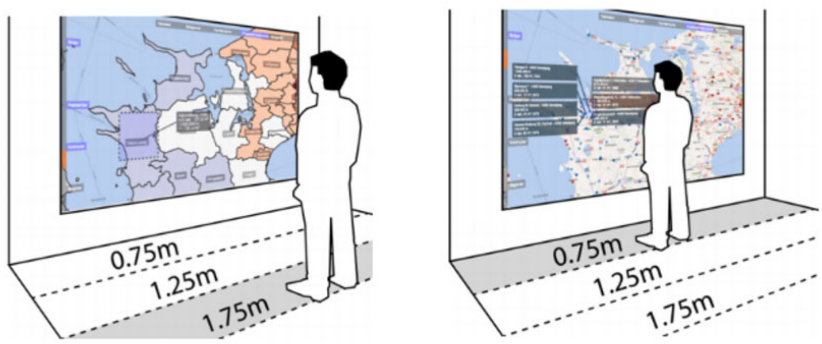


Figure 11: Using body proxemics (position) interaction to zoom in on content. Picture from [Jakobsen et al., 2013].

A Go-go interaction method for virtual interfaces is described by [Poupyrev et al., 1996] who implemented a non-linear mapping of an arm extension to reach objects normally out of reach. The technique was inspired by the cartoon character, Inspector Go-Go Gadget (Figure 12), who has the ability to extend his arms beyond a standard reach.

To implement the Go-Go method, a local area around the user is defined and in this area the motion of the user's virtual hand is mapped directly to their physical hand. When the user extends their hands beyond the defined area, the virtual hand moves non-linearly apart from the physical hand, allowing a farther reach while preserving the hand grabbing metaphor. Their usability test shows that the Go-go interaction technique was intuitive and easy to use, as it followed the logical extension of a normal action (stretching of the hand).

[Cheng and Takatsuka, 2009a] investigated various methods for free hand interactions with large screen displays. They found that the full arm stretch was the most common pointing strategy, while the most accurate strategy

Figure 12: Inspector Gadget. Picture from [Mesterius, 2013].

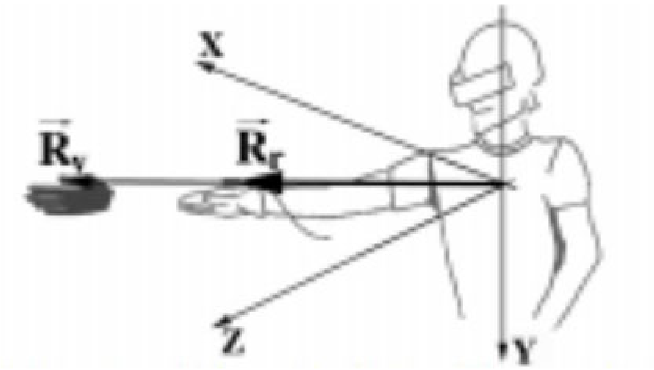


Figure 13: The go-go interaction method. The position of the real hand is defined by vector Rr, with the position of the virtual hand defined by vector Rv. Picture from [Poupyrev et al., 1996].

was when users lined up the target with their eye and fingertip. Figure 14 shows the various ways to implement a pointing vector.

[Ackad, Kay, et al., 2014] found that when designing interactive display interfaces it is important that gestures are not overly complex. Gestures should preferably require the use of only one hand [Ackad, Wasinger, et al., 2013]. This allows users to switch between hands, reducing the impact of fatigue [Annett and Bischof, 2013] as well as allowing for situations where the user is holding an object like a drink or mobile phone in their hands.

2.5 SOCIAL GESTURES

The nature of interaction between the user and a gesture interface has also been subject to research. While some users showed a concern for social embarrassment [Brignull and Rogers, 2003], most interactions showed a bias towards playfulness and performance [Tomitsch et al., 2014]. This can be disadvantages if little attention is directed towards the system's content [Grace

Figure 14: Pointing strategies for interactive large screen displays. Picture from [Cheng and Takatsuka, 2009a].

et al., 2013], although by utilizing this play theory we can design content that can aid a system's learnability.

2.6 THE APPLICATIONS OF GESTURE INTERFACES

As part of our background research, we review the current gesture systems already available. There is especially a trend for the use of gesture interfaces in three main areas: retail, health and fitness, and in the public environment.

2.6.1 Retail

A common use case for gesture interfaces is in shopping and commercial applications. By overlaying virtual objects onto the real world, these interfaces provide an Augmented Reality (AR) experience which can be used to showcase products. Augmented Reality involves the fusion of virtual overlays into the real world. By supplementing virtual sensory cues, the technology can alter a user's perception of reality [HITLab NZ, 2011].

For their Kinect for Windows showcase, Microsoft presented a retail scenario where stores are equipped with large screen displays and Kinect sensors. These displays have an augmented mirror background where users can overlay virtual clothes onto their body to check out how they look. Users can navigate the system using hand/arm gestures as well as by voice recognition [Microsoft, 2013].

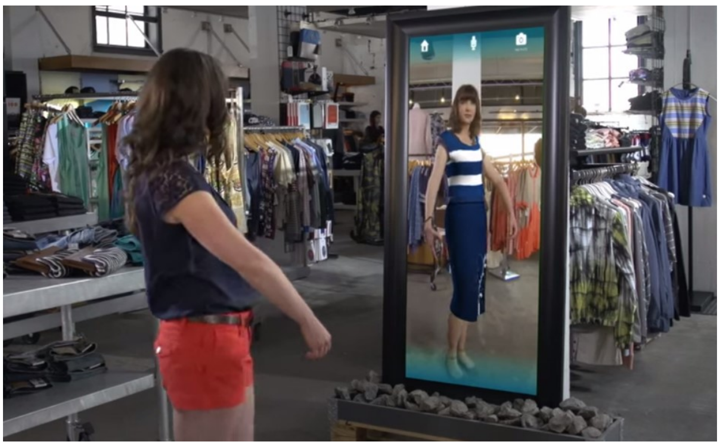


Figure 15: Microsoft's Kinect for Windows retail showcase. Picture from [Microsoft, 2013].

A similar setup retail setup is utilised by AR Door, a Russian company based in Moscow, who made a virtual fitting room for customers at a clothing retailer (Topshop). Using a Microsoft Kinect depth sensor, the system had an augmented mirror interface allowed customers to try on clothing virtually [ARDOORMoscow, 2011].

Figure 16: AR Door's virtual fitting room for Topshop, Moscow. Picture from [ARDOORMoscow, 2011].

The Puma store at Harajuku (Japan) has an installation of an interactive mirror with a "Virtual Fitting" function allowing users to check the look and styling of clothes. Users are able to interact with the system using touch gestures and take a photograph of themselves wearing outfits at different angles. Photos can then be shared with downloaded, printed or shared with friends on social media [INC, 2014].

Figure 17: Puma's interactive mirror installation in Harajuku, Japan. Picture from [INC, 2014].

Cisco introduced a similar concept of a virtual fashion mirror, called the StyleMe Virtual Fashion Mirror. By utilising gestures such as tapping or swiping their hands in mid-air, the system allows customers to browse their range of products and leave feedback, as well as taking photos to share it to their social network [Cisco, 2011].

Citibank's future branch concept utilises wall-sized interactive displays streaming current news, weather and financial updates. Customers are

Figure 18: Cisco's StyleMe Virtual Fashion Mirror. Picture from [Cisco, 2011].

shown walking up to the interactive wall and touching the screen to interact with the system [Citi, 2010].

Figure 19: Citibank's future bank concept. Picture from [Citi, 2010].

2.6.2 Health and Fitness

Another common use case for gesture technology is in the health and fitness sector. [Rydén et al., 2011] explored the use of a Kinect depth sensor to create a real time haptic fixture, able to guide a surgeon's hand during remote or robotic surgery. [Chang et al., 2011] demonstrated that the use of a Kinect based platform for patients with motor disabilities was a viable tool for rehabilitation. A commercial example is shown with the Nike+ Kinect Training gaming application [Rose, 2012]. In it, the user is represented by a virtual human avatar that follows the user's movements and body position. The objective of the application is to move your body following a workout routine as demonstrated by a 3D model on the screen. The user accumulates points that can then be uploaded online to compete with other users of the application.

Microsoft has also explored the use of gesture interfaces as a way for users to view, manipulate and control imaging reference in a surgical setting,

Figure 20: The Nike+ Kinect Training application. Picture from [Rose, 2012].

where traditional input devices might present a high risk of contamination [Microsoft, 2015f].

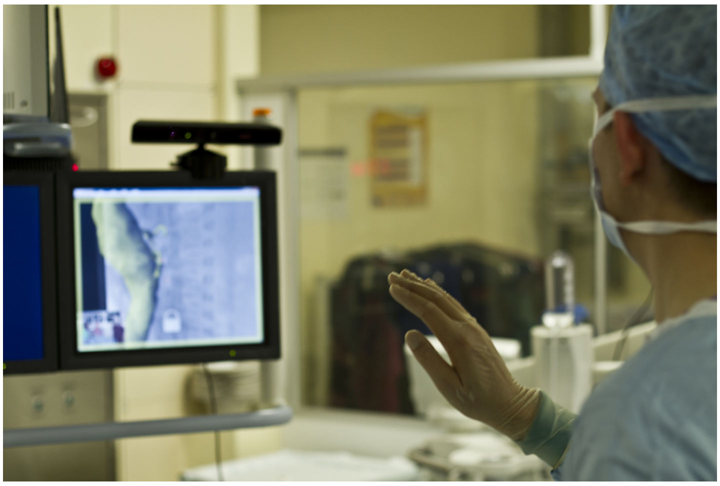


Figure 21: The use of gesture systems as a touchless interface. Picture from [Microsoft, 2015f].

2.6.3 Public Environments

The last common use case for gesture interfaces is as public displays; in places such as shopping malls, libraries, or airports. By adding gesture recognition capabilities, there is potential to make public displays interactive, providing a hub for information, entertainment or even community engagement [Peltonen et al., 2008]. A user interface without the need for additional equipment also allows the system to interact with a greater range of users than traditional computing, positioning the use of gestural interface as ideal for public interactive display. It does, however, presents its own issues. Due to the nature of being a public display, there is limited interaction time and therefore a tendency towards ‘casual’ users rather than ‘expert’ ones. Therefore, the usability of such systems is a priority when focusing on the public domain. Public interactive displays can provide a variety of information including map navigation, showcasing local business, upcoming events information, or even just for a fun experience. Accenture Technology's Interactive Network installation (Figure 16) comprises of an interactive large screen display at airports capable of displaying the latest news, weather, sports etc. Multiple users can simultaneously interact with the screen using touch gestures [Accenture, 2006].

Figure 22: Accenture Technology's Interactive Network interactive large screen display. Picture from [Accenture, 2006].

Another example is the Interactive Window made by Nuformer, consisting of a real-time video projected onto a street window (Figure 17). Passerbys are able to see an image of their reflection, augmented with special effects like a flaming outline or skeleton tracing [NuFormer, 2013].

Figure 23: Nuformer's Interactive Window. Picture from [NuFormer, 2013].

Interactive large screen displays are also actively used in corporate events. For the 2013 World Petroleum Congress, oil company Chevron had a large interactive wall where attendees were able to interact with videos and graphics on the screen. Content could be manipulated by the user's hand gestures, as well as changing content displayed based on the user's their proximity to the screen. The system also supports multiple user interactions [Control Group, 2015].

Last on our review, the National Museum of Nature and Science (Tokyo) presented an interactive public display for their Genso no Fushigi (The Wonder of Elements) exhibition. Their system enabled users to control an avatar,

Figure 24: Chevron's interactive wall. Picture from [Control Group, 2015].

an augmented character representation of their body silhouette, to interacts with a virtual world through their body/limb movements [IMGSRCinc, 2013].

Figure 25: The National Museum of Nature and Science's interactive public display, Tokyo. Picture from [IMGSRCinc, 2013].

2.7 A SUMMARY OF GESTURE INTERFACES

Our review of the current state of gesture technology has shown a rising use of gesture interfaces in applications such as retail, health and fitness and in the public domain. Interactive large screen displays have been used to provide retail customers with an enriched shopping experience. They have also proved useful in the medical field as a rehabilitation aid, or as hygienic way for controlling a system via touchless interaction (see Figure 21), and in public environments to increase community engagement.

Most gesture systems reviewed utilised similar interaction methods, using gestures pre-defined by researchers, hardware manufacturers or exporting those commonly used in existing touch interfaces. As such the usability, and suitability of such interactions is still not clearly defined, and is the focus of our research.

3 DESIGN CONSIDERATIONS

In the Chapter 2 we reviewed the current gesture systems available. We found that as gesture interfaces are typically put in a public or semi-public area, users are usually opportunistic and limited to a short interaction time, necessitating the need for an intuitive a learning curve and a low barrier of interaction.

3.1 USABILITY

Designing a system for a diverse range of user abilities present challenges and there are a variety of factors to consider. In his concept of Universal Usability, [Shneiderman, 2000] classed these into three categories:

- Extensive variations in equipment used- screen size, hardware performance, etc.
- User variability- range of user needs, ability and background.
- Learning curve of a system.

These aspects will be relevant in the development of our gesture interface. The social acceptance of gestures is also an important consideration that can impact a user's willingness to engage with an interface [Buerger, 2011]. The adoption of gesture technology can be hindered if it necessitates new behaviour that might be considered disruptive or embarrassing, as users often have to account for behaviour that can lead to embarrassment as well as attracting the attention of bystanders. [Rico and Brewster, 2010] suggest that it is the appearance of the gesture, rather than its energy efficiency, that influences it's acceptability.

3.2 LEARNING CURVE

The classification of gesture interfaces as a natural user interface comes with the promise of a low cognitive barrier of interaction. The idea of 'direct manipulation' was developed by [Shneiderman, 1993] to present a way of improving the user experience of a system. It can be broken down into four principles:

- The object of interest should always be present
- The use of physical actions are better than syntax
- Interactive objects should have an immediate response
- An incremental learning approach allows for users with minimal prior knowledge and presents opportunities for more complex interactions.

Shneiderman reasoned that the brain is hardwired for direct manipulation and by minimising the gap between a user's intention and the execution of a system task, user satisfaction can be increased.

Keeping the interaction threshold low with a simple interface flow is recommended to reduce the user's cognitive load [Müller, Alt, et al., 2010]. This can mean the sacrificing of unnecessary but interesting content. If complex techniques are required to complete a task, [Buerger, 2011] suggests that it be discoverable by experimentation or through a logical learning approach. The use of system tutorials are not an effective way to engage users, with most ignoring them and opting to learn an interface in real-time [Ackad, Kay, et al., 2014].

3.2.1 Immediate Usability

The term immediate usability describes an interface which a user is able to engage with spontaneously. This can be a system that provides a high motivation for a user to interact with (e.g. a fridge), one that presents a high attraction for users (e.g. a gaming console), or one that is easy to learn (e.g. a light switch). With any system there is always a gap between a user's knowledge and the interface. Designing for immediate usability aims to reduce this gap and presents an important consideration when designing our interface. [Kules et al., 2003] provide an example of immediate usability in the design of automated teller machines (ATMs), a device that has undergone extensive usability testing to provide a seamless user experience. Combined with a high motivation of users to use them (to get money), the usability of ATMs allows users with a range of abilities and background to interact seamlessly without prior experience of the system.

Another way to reduce the learning gap is by combining elements from one environment into another, i.e. to augment reality. By providing a representation of the user (e.g. a mirror image or silhouette, see Figure 27) in a gesture interface, we present an extension of the user into the virtual environment. This 'embodiment' interaction has been shown to increase a user's cognitive stimulation and perceived control of a system, allowing interaction with a system feel more "natural" [Lee et al., 2012].

Avoiding the need for user calibration or the use of an initialisation sequence (like a guided tour) can also reduce the interaction barrier a system, although it may impact on the subsequent experience if an interface has a complex learning curve.

3.2.2 Affordance

The term affordance was introduced by Donald Norman in his book, The Design of Everyday Things [D. A. Norman, 2013]. It describes the relationship between a physical object and an interacting user. The properties of that object and how it can be manipulated by the interactor is its affordance. Therefore, the affordance of an object refers to the possible 'interaction qualities' of an object and can be used to suggest the capabilities of a system. An example is shown with the buttons in (Figure 26). The button on the left has a shadow, giving it an apparent 3D look. The gradient shading also suggests a rounded surface. This skeuomorphic design takes properties from a physical button, giving the button an affordance of appearing "pushable."

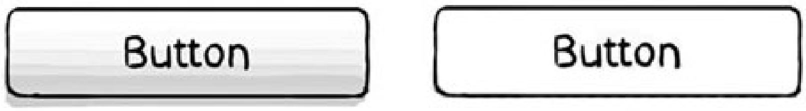


Figure 26: The button on the left has an affordance design to appear "pushable" when compared with the button on the right. Picture from [Demodern, 2014].

3.2.3 Display Blindness

Currently a large majority of large screen displays are passive. This presents a challenge for gestural interfaces, with the need to reveal themselves as interactive. This problem has been termed "display blindness"[Müller, Wilmsmann, et al., 2009]. One method to overcome this display blindness is the use of a dynamic representation of the user. This can be in the form of a mirror image, silhouette, or even a skeletal pictogram on the interface; providing an affordance that a system is interactive and has been shown to increase user engagement [Tomitsch et al., 2014]. Examples of dynamic user representations are shown in (Figure 27). One limitation of using dynamic representation is the risk of distracting user's from primary goal of the system, and more work is needed to evaluate this [Grace et al., 2013].

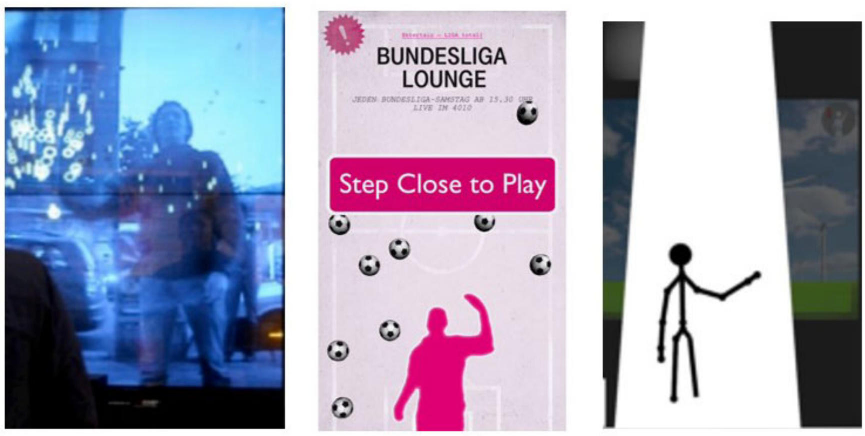


Figure 27: Dynamic user representations on gesture interfaces; (left) mirror image [Müller, Alt, et al., 2010], (center) silhouette [Müller, Walter, et al., 2012], (right) skeletal avatar [Grace et al., 2013].

In their Magical Mirror installation [Michelis and Müller, 2011] developed an interactive mirror display augmenting various optical effects like a glowing aura around the user's silhouette and providing various trail effects following moving objects on the screen. As the augmented mirror reacted to the user's body movement, it proved an effective method to overcome display blindness [Michelis and Müller, 2011]. The use of a silhouette representation of passing users was also found to be effective [Müller, Walter, et al., 2012].

The Looking Glass installation by [Müller, Walter, et al., 2012] explored various levels of user augmentation, with a mirrored user image producing 90% more engagement compared to a non-mirror, standard interface.

A silhouette interface attracted 47% more engagement than a non-mirror interface.

The utilisation of a mirror-like interface presents positive affordance and usability properties for a user. Various commercial displays (see Chapter 2.6) have incorporated a mirror-like design, thus for our experimental prototype, we will create a gestural system with an augmented mirror interface.

3.3 THE PHYSICAL INTERACTION ZONE

The concept of a physical interaction zone (PhIZ) was proposed by Microsoft's Kinect development team [Microsoft, 2015b]. It consists of a spatial zone around the user in which there is optimal tracking and gesture ergonomics. The PhIZ is an area relative to the user, spanning from the user's head to mid abdomen, a region in front of the user with their hand approximately 60% extended of their full arm length. A diagram of the PhIZ is shown in Figure 28.

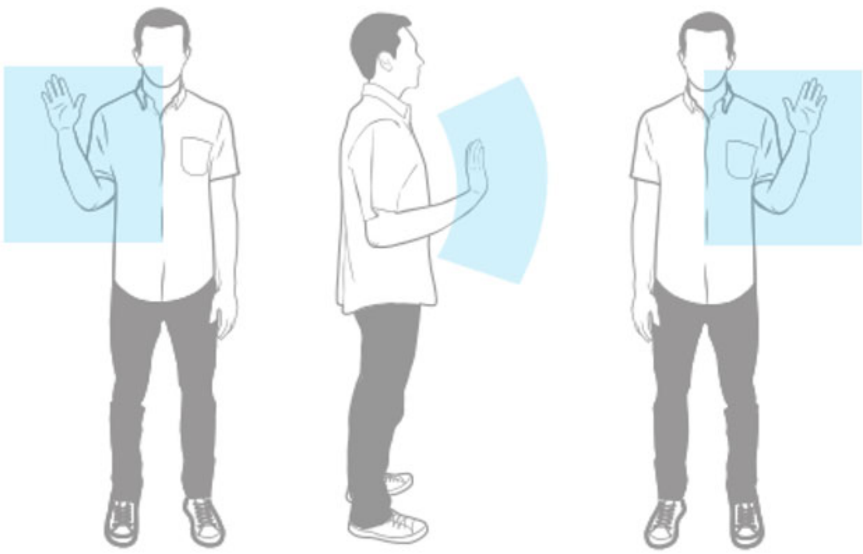


Figure 28: The Physical Interaction Zone. Picture from [Microsoft, 2015b].

The concept of an optimal interaction zone is support in the observations of users interacting with the Magic Mirror display by [Michelis and Müller, 2011], where users tend to position themselves into the centre of a display, naturally entering the 'interaction zone' in front (Figure 29).

3.4 DESIGN GUIDELINES

In Chapter 2 we learned about the gesture systems available and the various approaches available for the development a gesture interface. In this chapter we learned about ways to increase the usability of our system and how to effectively engage a user. Before discussing our prototype implementation, we summarise the main factors to consider for developing a successful gesture interface:

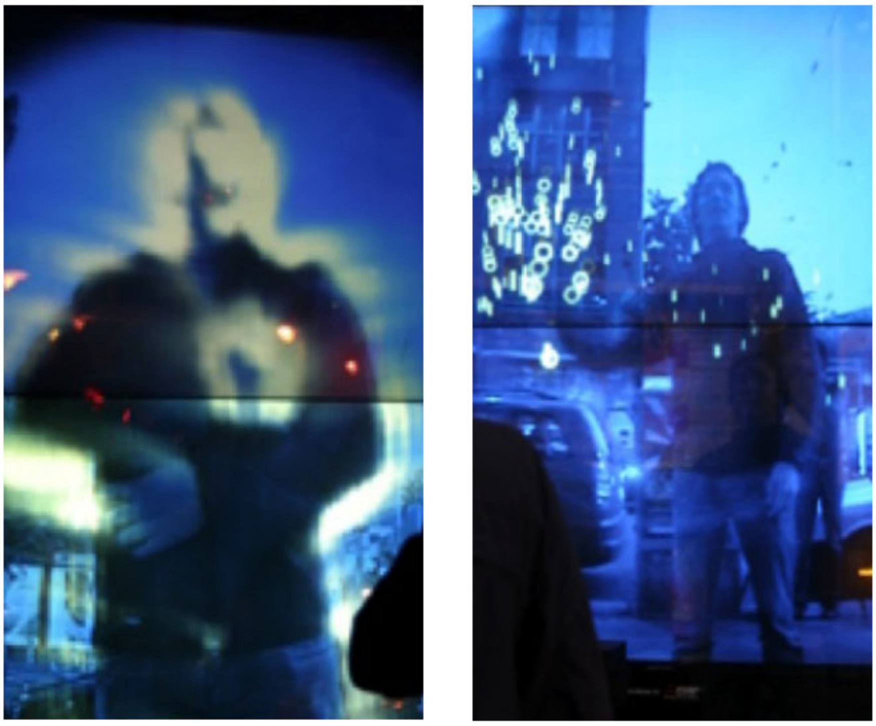


Figure 29: Users positioning oneself in the centre of a display. Picture from [Michelis and Müller, 2011].

- A low or incremental learning threshold will increase a system's usability
- Designing for affordance can aid the learning of a system's properties
- There is an optimal interaction zone for gestures
- Gestures should be able to be performed with one hand, and should be logical (e.g. a fast gesture should not be used to implement a slow task)
- Consider the social acceptance of gestures

[M. Nielsen et al., 2003] stresses that when developing a system, the objective should not be "to make a gesture interface," but to develop one in regards to a specific task and/or application. With this in mind, we discuss the hardware and software tools we used in the development of our prototype.

4 PROTOYPE DEVELOPMENT

There are various hardware are available to develop a gesture interface. The Microsoft Kinect depth sensor is one that is readily available and has been used in some of the commercial interfaces that we have reviewed in Chapter 2.6. For our research we used the Microsoft Kinect sensor running custom software made with the Unity development platform [Unity, 2015].

4.1 THE MICROSOFT KINECT SENSOR

The Kinect sensor is a motion-sensing device developed by Microsoft in 2009. The sensor is capable of capturing 3D data by emitting an array of beams from an infrared emitter, which are reflected by physical objects. The reflected beams are then captured by an infrared depth sensor that is able to calculate the distance of an object from the sensor. The sensor also includes a multi-array (4) microphone able to record localised audio and a tilt motor to move the sensors from -27 to +27 degrees from a horizontal level plane. A 3 axis accelerometer is also included in the unit, allowing it to detect the orientation and tilt of the sensor. A RGB (colour) camera is also included with a resolution of 1280x960 pixels at 30fps [Microsoft, 2015d].

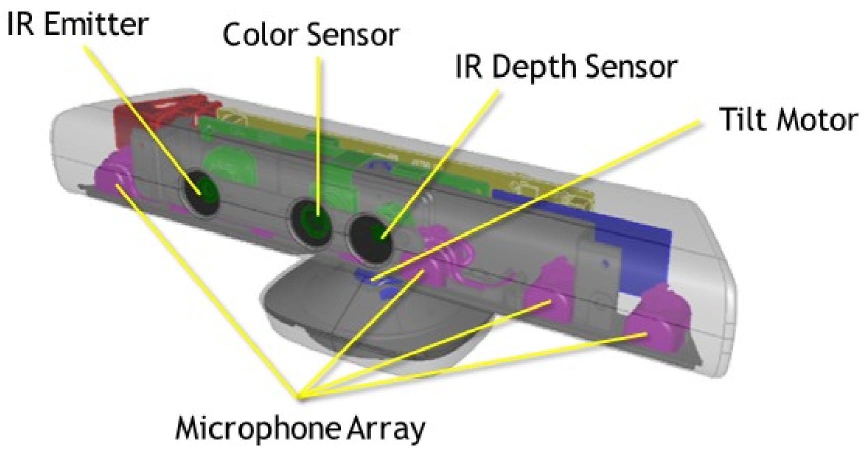


Figure 30: The Microsoft Kinect sensor (v1). Picture from [Microsoft, 2015d].

The sensor's range of detection is from 0.40 to 3 meters with a horizontal field of view of 57.5 degrees. The vertical field of view is 43.5 degrees, with an additional 27 degrees possible on both sides from the tilt motor. These are illustrated in (Figure 31) and (Figures 32).

A second generation Kinect (v2) was released in mid-2014, featuring additional sensors for increased tracking resolution. For this research however, we will be using the original Kinect (v1). This was due to the fact that the Kinect v1's ability was sufficient for our requirements, it was widely available and because the official Kinect v2 software development kit was yet to be released on commencement of this research.

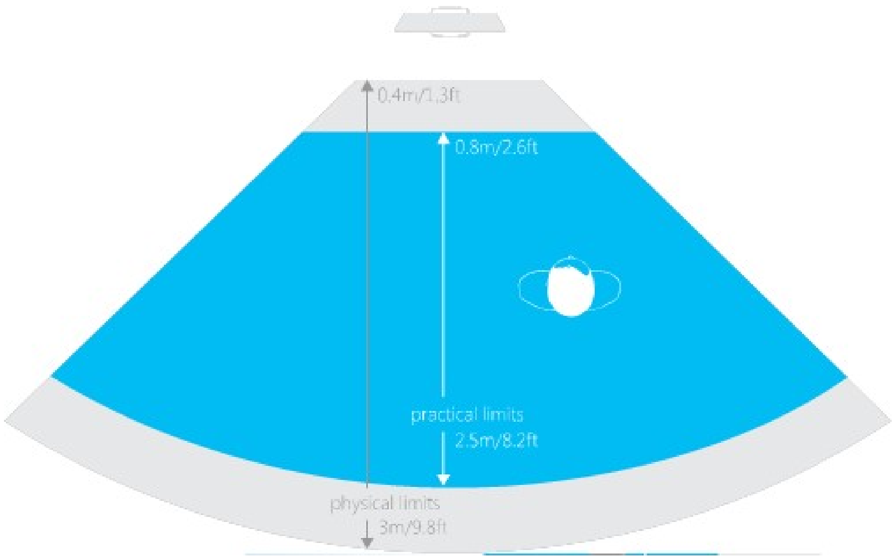


Figure 31: The Kinect's detection ranges from 0.4 to 3 meters, with 0.8 to 2.5 meters being the optimal distance. Picture from [Microsoft, 2015e].

4.2 THE KINECT SOFTWARE DEVELOPMENT KIT

The Kinect Software Development Kit (SDK) provides the tools and an application programming interface (API) to enable developers to create applications using the Kinect sensor. We will be using the 1.8 version which was published on 9/13/2013 and was the last SDK before the release of the second generation Kinect sensor [Microsoft, 2015c].

The Kinect SDK provides a skeletal tracking feature able to recognize and track up to 6 users in the sensor's field of view. Detailed tracking of up to two users is possible, with 20 skeletal joints to be tracked for each user (Figure 33).

The tracking algorithm is designed to recognise users facing the sensor in either a standing or sitting pose. The tracking of sideways poses is challenging, as part of the user is not visible for the sensor. No specific pose or calibration action needs to be taken for a user to be tracked.

While the Kinect sensor is capable of detecting audio, it is a feature that we did not use for our research. Development of a multi-modal display was considered (using gesture and voice) but the reliability of speech recognition is still lagging behind gesture recognition technology.

4.3 UNITY DEVELOPMENT SOFTWARE

To implement the prototype we chose to use the Unity game development software, version 4.64f1. Unity is a cross-platform development platform, with support for scripting languages like C#, UnityScript (also known as JavaScript for Unity) and Boo. The Unity software is also able to connect to the Kinect SDK using custom plugins and thus was well suited for the development of our gesture interface(Unity, 2015). A screenshot of the Unity development workflow is shown in Figure 34. Microsoft Visual Studio 2010 was also used to manage the Unity programming scripts.

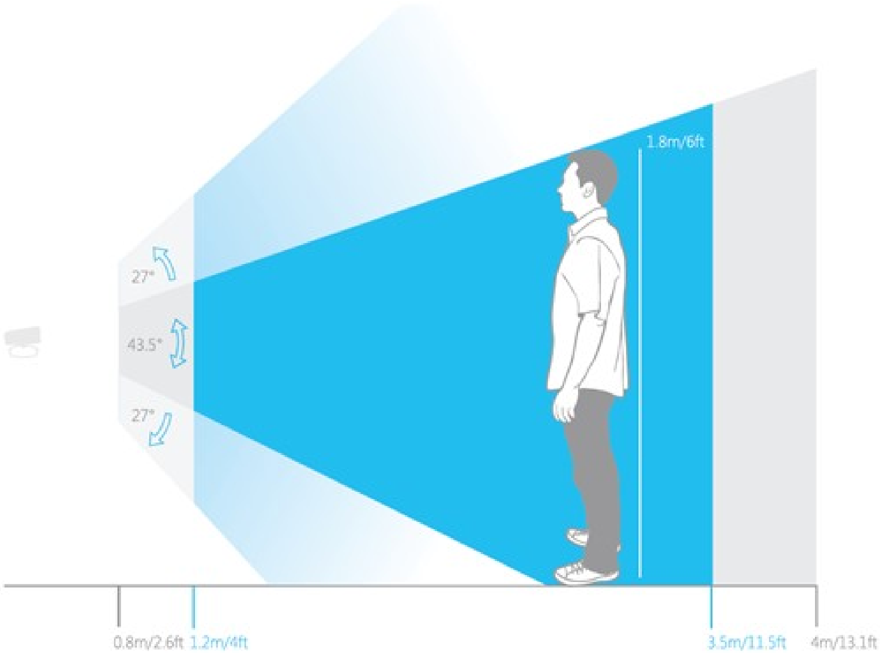


Figure 32: The Kinect's vertical field of view is 43.5 +/- 27 degrees with the tilt motor. Horizontal field of view is 57.5 degrees. Picture from [Microsoft, 2015e].

4.4 IMPLEMENTATION

In summary, we have reviewed the tools available for the development of a gesture interface. Our setup consists of a Kinect depth sensor, able to track a user's body movement and a custom graphical user interface built on the Unity development platform. The ability of the Kinect SDK, with its skeletal tracking feature, to integrate with the Unity development platform makes it an appropriate choice for rapid prototyping.

As we have learned in Chapter 2, various gestures can be used to execute similar actions, some more intuitive than others. To explore this, we design a guessability study, where users are able to define what gestures they feel are natural or appropriate for interacting with our interface.

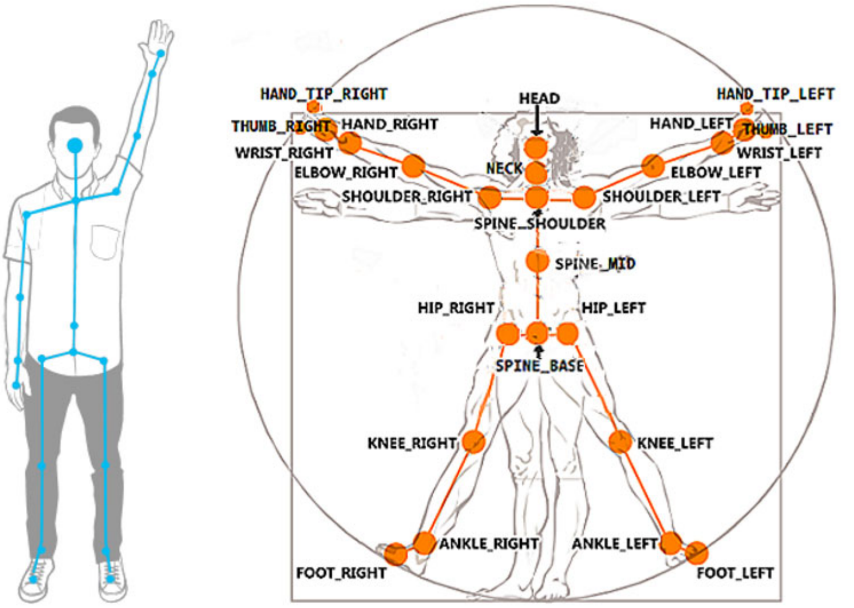


Figure 33: The Kinect SDK. Outline of the skeletal tracking (left), the joints tracked on the user (right). Picture from [Microsoft, 2015c].

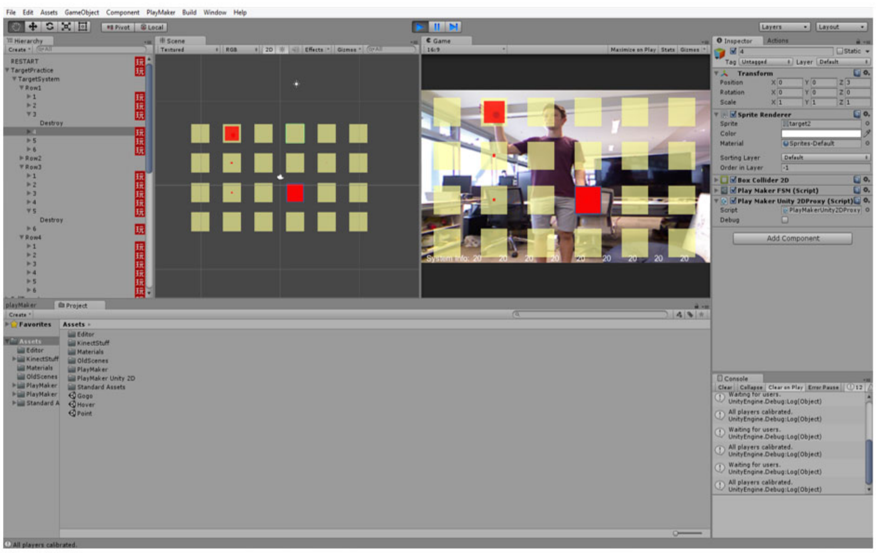


Figure 34: A screen-shot of the Unity game development platform.

5 THE GUESSABILITY STUDY

Following our background review, we decided to implement a guessability study to explore the gestures people naturally use for performing certain tasks when interacting with a gesture interface. A guessability study follows the style of a Wizard of OZ experiment, which is where a participant interacts with a computer system that they believe to be autonomous, but which is actually simulated by the experimenter. Wizard of Oz experiments have been shown to be an effective way to develop a user-centred gesture vocabulary [M. Nielsen et al., 2003] and have been used in the development of touch surface interfaces [Wobbrock, Morris, et al., 2009], mobile devices [Ruiz et al., 2011] and Augmented Reality applications [Piumsomboon et al., 2013].

5.1 THE SETUP

The experiment is held in an empty meeting room with a hardware system setup as shown in Figure 35. The system consists of a 52 inch Full HD television screen set up in a portrait orientation, running a custom application created with the Unity game development platform. There is a camera mounted on the left side of the screen providing a background camera stream to simulate a mirror. A Kinect sensor mounted on top of the display is used to capture body movement data.

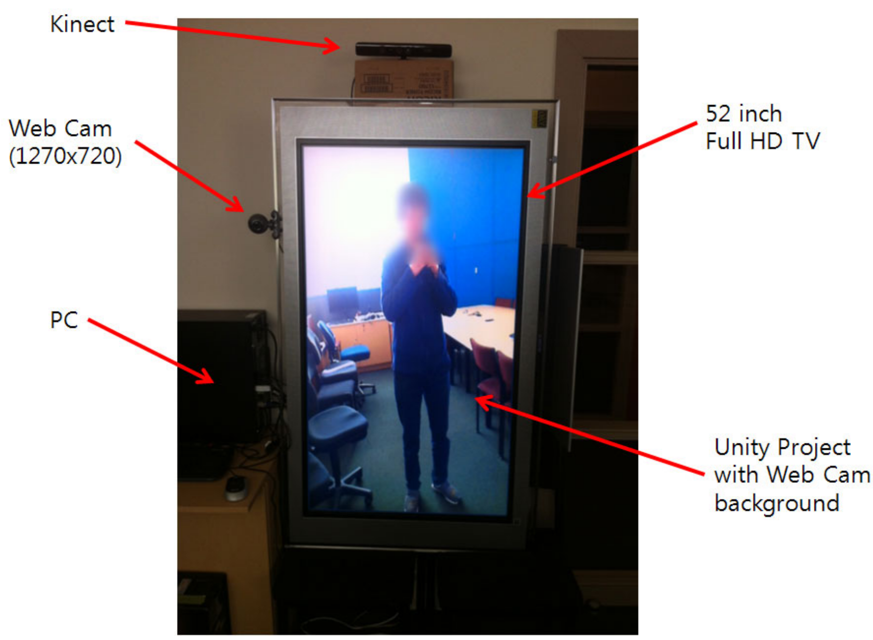


Figure 35: The experimental setup of our guessability study.

To record the movements of the participant, the experimental setup has an additional observation camera beside the participant to record both the participant's gesture and the animation on the screen. The Kinect sensor data, observation application and the visualisation software are run on three separate computers. The software for our guessability study consists of a range of animation scenes overlaid on a live video background, streamed from the camera. The animations visualised on the screen demonstrates the tasks users are required to perform using gesture interaction.

5.2 THE TASKS

For this guessability study, the aim is for the user to propose a gesture they feel appropriate to perform for common tasks encountered when interacting with a gesture interface. We identified six common tasks that may be useful when interacting with such systems. An animation sequence is made for each task, simulating what would occur on the interface when a user is using the system to complete the specified task. The six tasks are summarised in Table 1:

Task	Description	Analogy
Pointing	Point at various objects on the screen	Moving a mouse cursor
Selection	Select an object on the screen	Mouse click
Drag (and drop)	Drag an object and move it around	Mouse drag/drop
Dichotomous option	Choose between two fixed options	e.g. YES/NO option
Horizontal scroll	Scroll through a horizontal list of objects	e.g. Scrolling a web page
Vertical scroll	Scroll through a vertical list of objects	e.g. Scrolling a web page

Table 1: The tasks in the guessability study.

With [Tomitsch et al., 2014] showing a tendency of gesture interface users to show a 'playfulness' attitude, the user interface animations and graphics in our system were created with a gaming style. Screenshots of each task are shown in (Figure 36. The background texture of each scene presents as an augmented virtual mirror visualisation, which can be turned off (appearing black as in the screenshot).

5.3 THE QUESTIONNAIRE

For the guessability study we also included a questionnaire section broken down into three components. The first component included general questions to gather information into the participant's demographic and experience with gesture interface. The second component evaluated task usability on a 7 point Likert Scale, a commonly used research method for measuring system usability [Brooke, 1996]. An example is shown in Figure 37. The last component of the questionnaire provided an overview of the whole experiment, allowing for comments and clarification of any issues that may have arisen. The questionnaire form is shown in Appendix A.3.

5.4 THE PROCEDURE

In this section, the standard procedure for the guessability study is described. Participants are first introduced to the objective and overview of the study- to explore user-defined gestures when interacting with a gesture interface. Participants are then given a consent form to read and sign (Appendix A.2). The participant then answers a pre-experiment questionnaire (Appendix A.3).

For each of the six task (described in section 5.2):

1. The participant is shown an animation depicting the execution of a task. In addition, the experimenter verbally explains the task in a short description while the participant watches the animation. The experimenter's run sheet and dialogue can be found in (Appendix A.1).

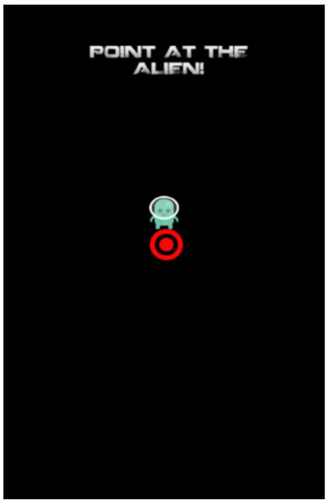
2. The participant is then asked to describe a gesture appropriate to complete the shown task.

3. The animation is replayed while the participant acts out their proposed gesture to simulate performance of the task. This step is repeated 3 times for each interaction task in order to let the participants become comfortable with acting out the gesture.

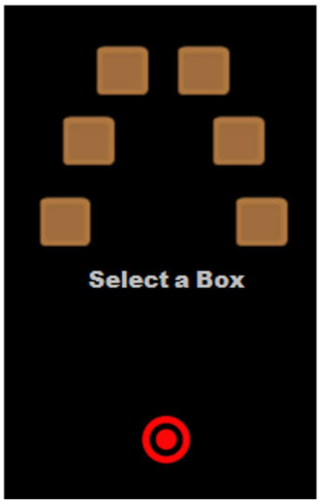
4. The participant then answers a usability questionnaire on the acted out gesture for the given interaction task. The questionnaire can be found in (Appendix A.3).

After all six tasks are completed, the participant is then required to answer a post-experiment questionnaire (Appendix A.3) and is debriefed to clarify and get more details on the questionnaire response. The order of the interaction tasks presented to the participant is randomised to counter balance the bias from the learning effect from fixed ordering.

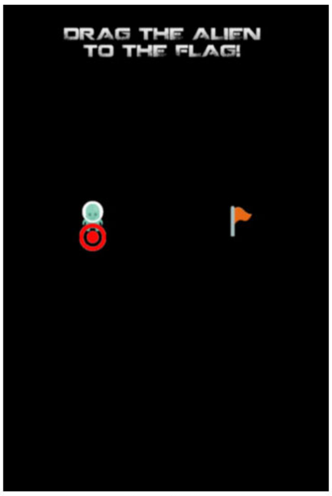
At the conclusion of the study we analyse the recorded gestures to identify intuitive gestures for each tasks, by investigating consensus among the participants.



Pointing

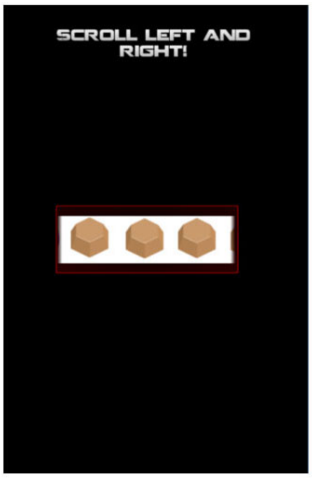


Selection

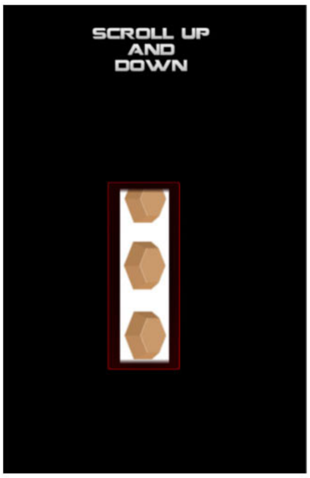


Drag (and drop)

Dichotomous option



Horizontal scroll



Vertical scroll

Figure 36: Screenshots of the task animations for the guessability study.

1. I was able to **perform the gesture well**.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

Figure 37: An example of a 7 point Likert scale. The average is set at 4.

6 RESULTS FROM THE GUESSABILITY STUDY

Once the results are collected from the user study, we analyse the collected data using qualitative and quantitative analysis methods.

The video recording is analysed by coding the type of gestures into categories. Once categorised into their respective types, we analyse the agreement score defined in [Wobbrock, Aung, et al., 2005] to show the level of agreement between the participants. This method had been applied in prior guessability studies [Piumsomboon et al., 2013; Ruiz et al., 2011; Wobbrock, Morris, et al., 2009].

Results from the questionnaires are mainly analysed with quantitative analysis on the Likert-scale rating questions. Mainly descriptive statistics are used for summarising the results, while for comparing between participant groups we use inferential statistics such as Chi-square tests or Mann-Whitney U tests.

The responses to the open questions are analysed through keyword extraction and coding. The results are summarised using descriptive statistics, complemented with quotations from the participant's responses.

6.1 PARTICIPANT ANALYSIS

We recruited 20 participants in total with a mixture of different cultural/technical backgrounds and gender. From the participants, 16 of them participated with the augmented mirror style visualisation, i.e. the background utilised a live video stream mimicking a virtual mirror. The remaining 4 participants performed the study with the same setup, with the exception of a non-augmented mirror visualisation, i.e. a black background. This was done to explore the effect of an Augmented Reality (AR) vs non-AR condition.

For the group with using the augmented mirror style visualisation, 9 of the participants were male and 7 females with ages ranging from 21 to 34 years old (Mean = 26.6). Participants had various cultural backgrounds with only four of them identifying English as their first language while the rest varied including Spanish, Indian, Chinese, Korean, Hokkien, German, French, Croatian and Arabic. However, all of the participants were fluent enough with English for participating in the study.

When asked about their previous experience using hand gesture interfaces (e.g. Kinect based games, Nintendo Wii), only three participants replied that they have used it "more than once a month" while the rest answered" a couple of times a year" or "not at all". More than half of the participants (9) had never used an augmented reality (AR) interface before, while only four of the participants had used it more than once a month. Most participants (9) stated that they used their right hand mostly for pointing and making gestures, while two state left hand, and five both hands. Rating on a 7-point Likert scale (see Figure 4.3) showed participants moderately agreed that they consider themselves using gestures a lot in everyday life (Median = 4.5, Range = [2 7]).

In the other group that used non-augmented mirror style visualisation, participant's ages ranged from 23 to 45 (M = 29.3), of which 3 were males.

6.2 TASK ANALYSIS

Here we summarise our analysis of the video recordings identifying gestures types, the level of agreement between participants, and metrics from our usability questionnaires.

6.2.1 The Pointing Task

For the pointing task, the user has to point at various objects on the screen, analogous to moving a mouse cursor. Participants mainly suggested two types of gestures: directing and hovering. With directing gestures, participants pointed their hand or finger at the target on screen. With hovering gestures they placed their hand (or finger) to align their mirror image (in the video stream) to the target position. The difference is shown in Figure 38. Twelve (80%) participants used the hovering method while three (20%) used the directing method, which is significantly different from random choice ($X^2(1) = 5.40$, $p = .020$). The Level of Agreement (A; calculated as a squared sum of proportion of each category) of 0.68 shows there is good consensus among participants. As pointing is also one of the common sub-tasks for selection and dragging, we analysed those instances and obtained similar results as summarised in Table 2.

Main Task	Hovering Gesture	Directing Gesture	Other Gesture	Level of Agreement
Point1	12	3	0	0.680
Select	12	3	1	0.601
Drag	13	3	0	0.695

Table 2: Types of pointing gestures.

Figure 38: For pointing tasks, participants used a hovering gesture (left) or directing gesture (right). Note that the participant is crouching out of sight in the hovering photo.

The participant's choice of using a finger or open hand for the pointing tasks varied according to context (see Table 3). While participants tended to point with their fingers more (67%) when pointing is the main task, when pointing is a sub-task of selection or dragging they preferred using an open hand (53% for selection and 69% for dragging, $X^2(1) = 1.22$, $p = .269$ and $X^2(1) = 3.89$, $p = .049$, respectively).

Main Task	Gesture	Open Hand	Pointing Finger	Level of Agreement
Point	Hover	5	7	0.514
	Direct	0	3	1.000
	Subtotal	5	10	0.556
Select	Hover	7	5	0.514
	Direct	1	2	0.556
	Subtotal	8	7	0.502
Drag	Hover	9	4	0.574
	Direct	2	1	0.556
	Subtotal	11	5	0.570

Table 3: Type of hand poses in pointing gestures.

6.2.2 The Selection and Drag (and Drop) Tasks

For the selection task, the user has to select a graphical object on the screen. For the drag (and drop) task, the user has to first select an object, triggering a draggable state, then release the object, analogous to the drag and drop action when using a mouse. These tasks are summarised together for their similarities.

For selecting objects, participants suggested three approaches: waiting, tapping, and grabbing. A waiting gesture involved holding the cursor on an item of interest for couple of seconds. For the tapping gesture, users moved their hand (or finger) towards the screen in a pushing motion. The grabbing gesture involved the folding of the fingers to close hand, forming a fist.

With selection as a main task more than half of the participants performed tapping gestures, with the grabbing gesture the next most common alternative. In comparison, with dragging as the main task, selection was achieved predominantly using a grabbing gesture. The results are summarised in Table 4.

Main Task	Waiting	Tapping	Grabbing	Level of Agreement
Select	2	9	5	0.430
Drag	4	3	9	0.414

Table 4: Types of triggering gestures for the select and drag tasks.

6.2.3 The Dichotomous Option Task

For the task of choosing between dichotomous options, the user has to choose between two fixed options, for example choosing from either a YES or NO option on the screen.

Most participants used one of two gestures: pointing directly at the options on the screen or making a thumbs up/down gestures at the desired option. Others gestures included nodding (for yes) or shaking their head (for no), raising hands (yes), and crossing arms (no). Some of these gestures are shown in Figure 39.

6.2.4 The Horizontal/Vertical Scrolling Tasks

For the scrolling tasks, users had to scroll through a horizontal and vertical list of objects.

Answer	Thumbs Up/Down	Direct	Other	Level of Agreement
Yes	7	6	3	0.367
No	6	6	4	0.344

Table 5: Type of gestures to choose between dichotomous options.

Figure 39: For the dichotomous option task, participants indicated yes and no by raising or crossing their arms (left) or making a thumbs up and down gesture (right), respectively.

A large number of users (81%) used swiping gestures resulting in high Level of Agreement score. While the majority of the users agreed that the swiping gesture is the most intuitive for scrolling task, there was variety of hand postures used while swiping, as well as variations in how they indicated the start and end of the swiping gesture. These are shown in Figure 40.

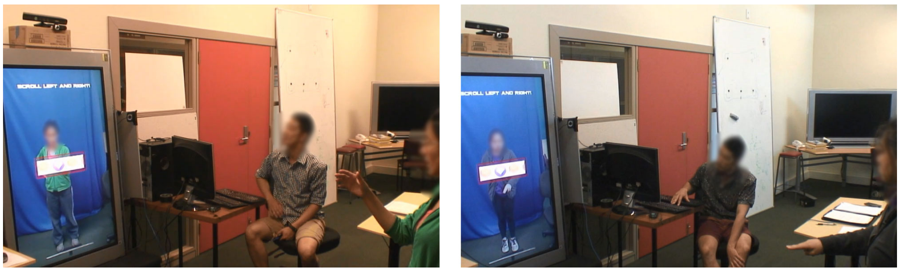


Figure 40: A majority of participants used swiping gestures the scrolling task. They either used an open hand swipe (left) or a finger swipe (right) gesture.

Many participants used an open hand for swiping while some swiped using one or two fingers. Likewise, some participants made a pressing gesture to start and end the scrolling, others used a hold and waiting method at either the start or end, while some even hid their hands behind their backs to signal the end of a gesture. While it was hard to determine the gestures used for indicating the start and end of swiping, we observed an interesting pattern that most of the participants positioned their hands on the graphical interface representing the list of items. All of the participants that performed a swiping gesture moved their hand over the graphical object list in the screen space, while even those performing other gestures placed their hand either on the list or near (slightly above or below) the list. A summary of the scrolling gestures used is found is shown in Table 6.

Swiping	Directing	Others	Level of Agreement
13	2	1	0.680

Table 6: Type of gestures used to scroll a list of items.

6.3 QUESTIONNAIRE ANALYSIS

The usability of user-defined gestures is evaluated and reported as followed.

When responding to the questions regarding the usability of the gestures performed, participants rated their answer on a 7-point Likert scale. Participants felt the gestures were well defined (Md = 5), intuitive (Md = 5.5), natural (Md = 5.5), easy to perform (Md = 6), and easy to learn (Md = 5). The results are summarised in Figure 41.

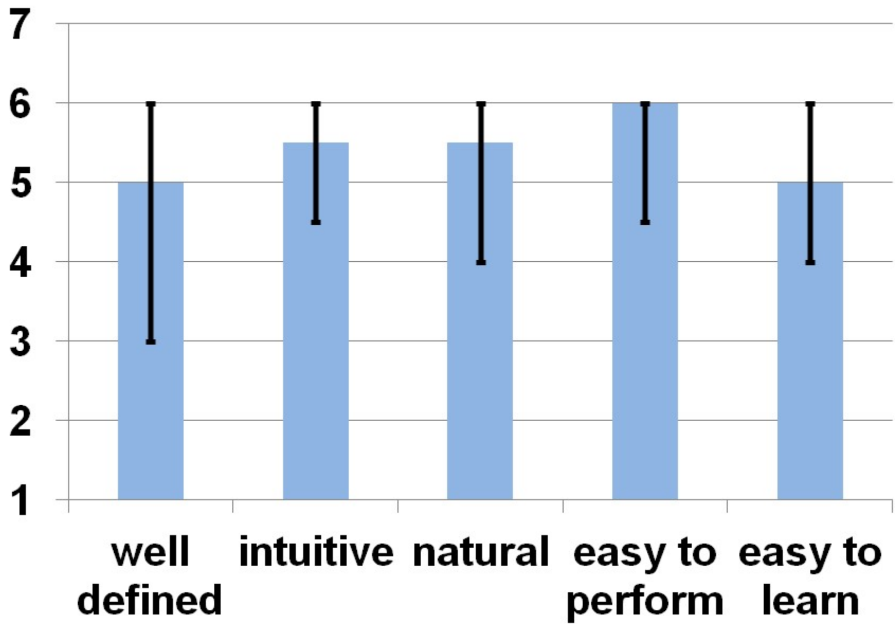


Figure 41: Results of Likert scale rating on usability. (1: not at all 7: extremely; whiskers represent inter-quartile range).

Participants also moderately agreed that gesture interaction would be comfortable to use in public spaces (Md = 5), and they would be more likely to use the public information display if using gestures (Md = 5) (See Figure 42).

When asked what kind of potential problems could arise when using gesture interfaces in a public setting, most participants (11 out of 16; 69%) worried that the system might fail to recognise or misinterpret their gestures. Other common issues mentioned included privacy (3), non-intuitive gesture definition (2), and accessibility for the impaired (2).

6.4 NON-AUGMENTED MIRROR VISUALISATION

As part of the design iteration process, we turned off the mirror visualisation effect to compare the difference in user-defined gestures, if any, when the interface had a lack of mirror feedback (displaying a solid black background

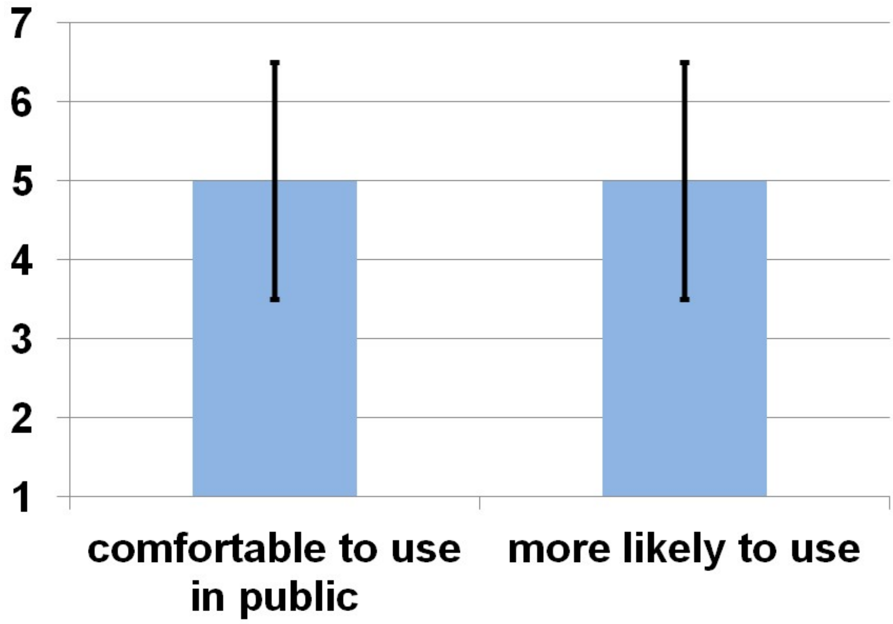


Figure 42: Results of Likert scale rating on overall usefulness of the system. (1: not at all 7: extremely; whiskers represent inter-quartile range).

instead). Four participants were tested with this condition (3 males; ages from 23 to 45, M = 29.3).

All four of the participants used directing gestures for pointing, implying that using the use of hovering as a deictic gesture may be unique to augmented mirrors. This could be due to the lack of sensory cues present as compared to the augmented mirror visualisation, which provided feedback on the user's hand position through the mirror image. However, the sample size was not big enough to draw statistically significant conclusions, so this should be investigated further in future studies.

6.5 SUMMARY OF THE GUESSABILITY STUDY

The results from our guessability study showed that the majority of participants used a hovering gesture for pointing tasks when interacting with our gesture interface, which had an AR mirror visualisation. When compared with non-AR displays, directing gestures were preferred. It was noted that for objects on the edges of the screen, or for those far from the user's reach, some participants did not bother to match their hovering positions exactly according to their image on the screen and instead only reaching as far as they could or felt like. This suggests that hovering based pointing gestures could be combined with other types of interaction techniques for reaching far targets. When comparing the mapping methods in pointing gestures, the level of agreement was low for hand poses in pointing gestures and triggering methods for selection and dragging gestures. One way to overcome this problem could be designing the system to support more than one type of gesture, allowing users to use their preferred method. The different poses are shown in Figure 43 (left column). For choosing between dichotomous options, participants used both deictic and metaphoric gestures. This suggests that metaphoric gestures could be optionally used instead of deictic gestures, reducing users' effort to point accurately. This is similar to how keyboard shortcuts work in 2D graphical user interfaces.

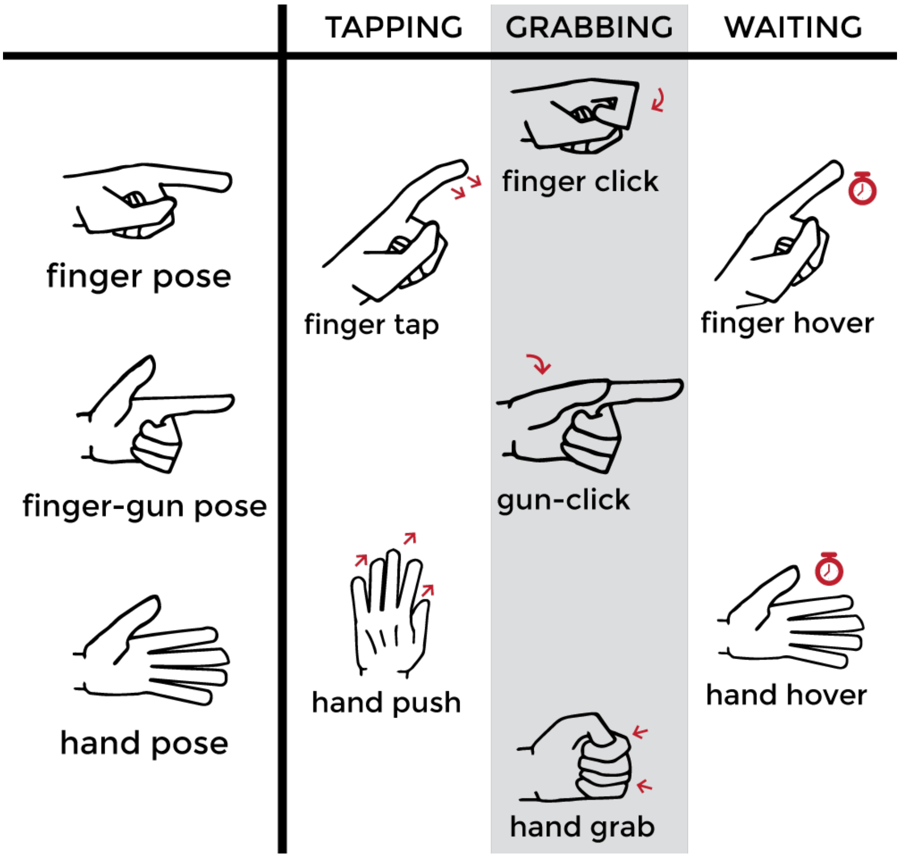


Figure 43: The various hand poses used for controlling the cursor position (pointing gestures) and the selection gestures used, categorise into the three main methods seen- tapping, grabbing and waiting.

6.6 NEXT STEP: FOCUSING ON ALIGNMENT

In this guessability study our results identified that for certain interaction tasks, there are gestures that a majority of users agree on, while for some tasks this is not the case. Hovering gestures were the most prominent gesture used for pointing when interacting with a gesture interface with an AR mirror style visualisation. Using the mirror visualisation as a reference, users aligned their hand position in the image to their targets. For targets that were far to reach, some of the participants did not bother aligning their hand positions exactly in the screen space, but just reached as far as they could (see Figure

As the target position of a user's focus, analogous to the position of a mouse cursor, is an essential function of most interface and presents as a subtask in other gestures, we decided to focus on evaluating further methods of positioning the interface cursor, specifically, to explore the usefulness of a non-linear mapping system of the cursor to a user's hand position.

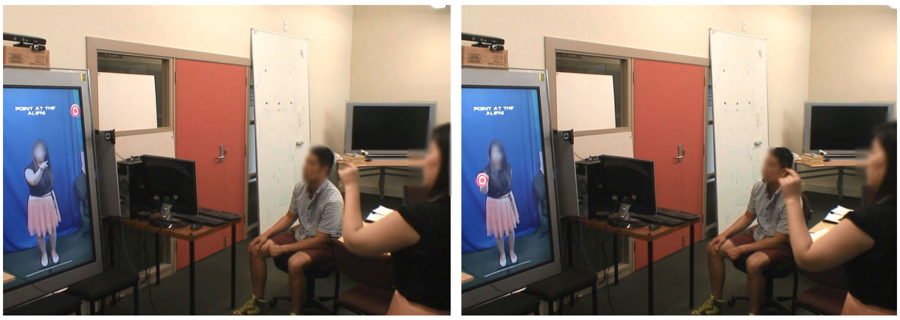


Figure 44: A participants using directing gestures for far targets (left), while using hovering gestures for close objects (right).

7 THE MAPPING STUDY

From our guessability study we found that users lack agreement in determining the most intuitive method of determining an interface's cursor position (the pointing task). While the hovering method was commonly used, users sometimes fail to directly position their hovering hand accurately when trying to reach far targets. To further explore this issue we conducted a user experiment to compare an alternative way to map the interface cursor, over the conventional one-to-one mapping (i.e. the hovering method).

We proposed a hybrid approach, combining a non-linear mapping between user's virtual hand position (cursor) with the hovering based pointing gesture in image space. This is similar to the go-go interaction method described in Chapter 2.4. The method we used for implementing this is shown in Figure 45. As the hovering method is interacting in the image space, we apply non-linear mapping in the image space by adjusting the calibrated projection parameters to the user's skeleton tracking information. H and S represent the position of the user's hand and shoulder in image space, respectively, while C is the position of the cursor following the user's hand. We first calculate the vector HS that originates from the shoulder and ends at the user's hand. While the norm of this vector $|HS|$ is less than a predefined threshold r, we use H as the position of C, which means the cursor is overlaid on the user's hand position. When $|HS|$ is greater than r, we scale the vector HS with the scaling factor s defined by:

$$s = \frac{|HS|}{r}$$

As a result, the position of the cursor C is calculated as:

$$C = S + sHS$$

This results in placing the cursor away from the user's hand, as if it is extended out towards the direction where the user is trying to reach. A screenshot of the working interface is shown in Figure 46(b).

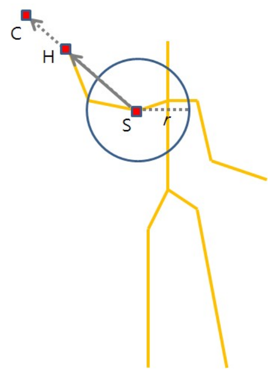
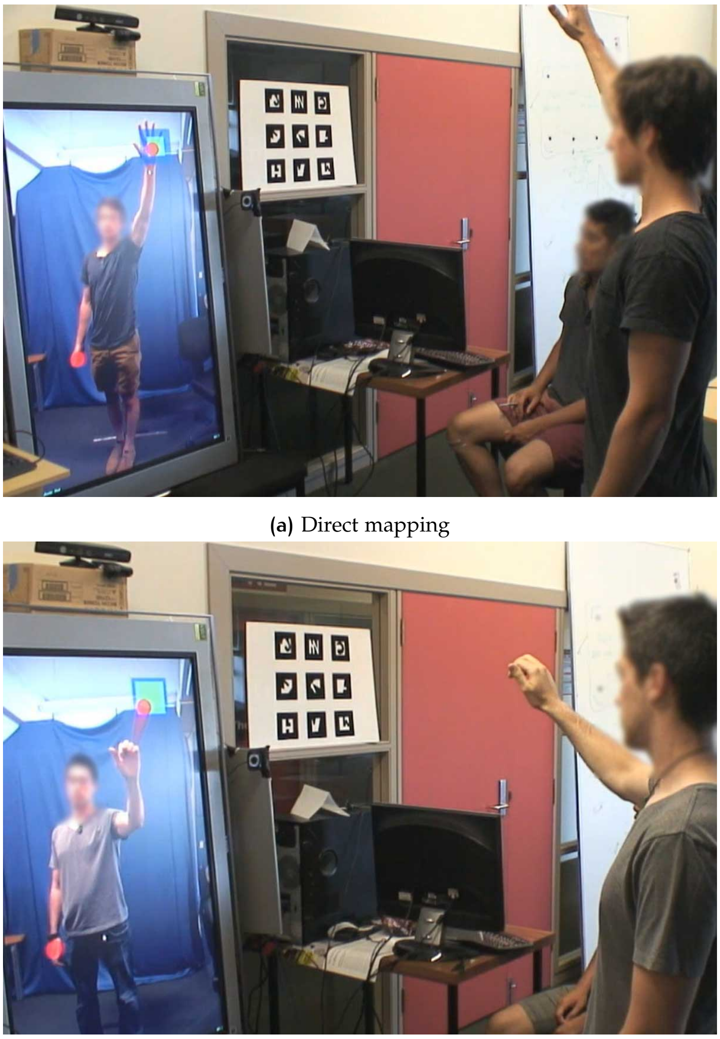


Figure 45: The non-linear mapping of the cursor to the user's hand position.

The region where direct mapping (i.e. the hand position is directly assigned to the cursor position) is applied can be defined in various ways. For instance, in our implementation we use a circular region centred on the shoulder, with a radius of the max length of upper arm in image space. We also filter out the region below the shoulder not to use non-linear mapping based on the observation that extended hand at the resting position feels unnatural, and also most of the objects below can be easily reached by the user without the help of the non-linear mapping technique. There could be also other approaches of defining the region for direct mapping using different geometric constraints defining various shapes and sizes of the region, and also applying heuristics depending on the target application.



(a) Direct mapping

(b) Non-linear mapping

Figure 46: The two interfaces in the mapping study.

7.1 THE SETUP

For the mapping study we used the identical hardware setup as described in Chapter 5.1 for the guessability study.

For the software aspect, we created a custom gesture tracking software that was able to track a user's hand position (using the Kinect depth sensor) and map the co-ordinates to a cursor position on the interface. The user's hand position could either relate to a cursor by direct mapping (for a hovering gesture), or via a non-linear mapping between the user's hand and cursor position. The software is also able to show a target box in random positions on the screen and had a timer to record the task completion time.

Participants are asked to stand in front of the gesture mapping interface at a distance where, when stretching their arms, they can reach the far corners on the screen (in the mirror image). While the system is capable of tracking and interacting with both of the user's hands, we asked the participants to use only their dominant hand for completing the experimental task, to assure the results are not affected by using different hands.

7.2 THE TASK

For this mapping study, the aim is for the participants to point at the blue target boxes on the screen using their dominant hand. A red circle on the screen represented the cursor, and could either be mapped directly onto the user's hand, which we will now call the *hovering* interface, or in a non-linear fashion, which we will now call the *extension* interface. The two interfaces are shown in Figure 46.

Once the cursor comes into contact with and stays on the target box for 2 second the target disappears and another target appears on the screen. While the participant is pointing on a target, an animation of a growing green box inside the blue target box is given to indicate the amount of time left until the target is selected and disappears.

No two consecutive targets were shown at the same position on the screen, forcing the participant to move their hand to the next target when it appears. The positions of the target are randomised between the participants but not between the conditions to make the two conditions balanced for a single participant. There were 27 targets to select in a single trial. Once the participant finishes selecting all of the targets a thank you message appears on the screen.

7.3 METRICS RECORDED

The following information is collected during the experiment.

1. Participants' response to the pre-experiment questionnaire, which includes demographic and participants' background information.
2. For each condition:
    a) Measured task completion time
    b) Subjective rating on various usability aspects of the given interaction method
    c) Written feedback or comment on the given interaction method
3. Participant's response to the post-experiment questionnaire including:
    a) Preference between the two methods

b) Written feedback on advantage and shortcomings of each method

c) Written feedback on overall experiment

The questionnaires for collecting subjective feedback from the participants consists of a number of Likert scale questions for rating usability and user experience, and also open questions to collect qualitative feedback. The questionnaires used for the experiment are shown in (Appendix A.5).

7.4 THE PROCEDURE

Participants are first introduced to the objective and overview of the study-to explore the different mapping conditions of pointing gestures. Participants are then given a consent form to read and sign (Appendix A.4). The participant then answers a pre-experiment questionnaire for collecting demographic information (Appendix A.5). The participant is then given an experimental task to do under two conditions (hovering interface vs extension interface). For each condition:

1. The participant is initially given a chance to interact freely with the system to get used to the conditions.
2. Once familiarised, the participant starts the task of selecting 27 targets on the screen.
3. After the task, the participant completes a per-task questionnaire.
4. The task is repeated under other condition.

Once both conditions have been tested, the participant answers the post-experimental questionnaire and gets a debriefing session to clarify any issues in the questionnaire responses. The order of the two experimental conditions is counter balanced by alternating the order of which interaction method is tried.

8 RESULTS FROM THE MAPPING STUDY

From the mapping study we collected a mixture of quantitative and qualitative data. For analysing the results of task completion time between the mapping conditions, we used a paired t-test. For comparing the results of Likert scale ratings from our questionnaires, we used the Wilcoxson Signed Rank tests, with the Chi-square test for evaluating the users' preference. The alpha level is set to 0.05 for all of these tests.

8.1 PARTICIPANT ANALYSIS

We recruited 10 participants for the extension experiment. All of them were graduate students, none of them were female, and their age ranged from 21 to 27 years old (M = 23.9). Participants used in the mapping study had not been previously involved in the guessability experiment. The lack of diversity among participants was due to a combination of factors- we had a tight schedule for conducting the mapping study, which coincided with the Christmas and New Years holiday period, leading to recruitment issues.

Only one of the participants answered that he uses both hands for making gestures in everyday life, while the rest of them answered that they predominantly use their right hand. When asked if they had used motion gesture interface before, three of the participants answered they have not used it before at all, two answered they use every day, while the answer of the rest varied from couple of times a year to couple of times a week.

8.2 TASK COMPLETION ANALYSIS

In terms of task completion time, participants spent about 15% less time with the hovering interface compared to the extension interface ($t(9) = -3.79$, $p = .004$). With the hovering interface, participants spent about 84 seconds in average (Std. Dev. = 4.5), while with the extension interface they spent about 99 seconds (SD = 11.9). The results are summarised in Figure 47.

8.3 QUESTIONNAIRE ANALYSIS

Results from the Likert scale rating questions showed both of the methods having good usability, as most of the aspects of usability were rated above the average (i.e. greater than 4 out of 7). While the hovering interface appeared to be marginally better in many aspects in comparison with the extension interface, participants felt the extension interface was slightly less physically stressful. A summary of the questionnaire results is shown in (Table 7).

When the participants were asked which method they preferred, seven participants (70%) answered they prefer the hovering interface, while the

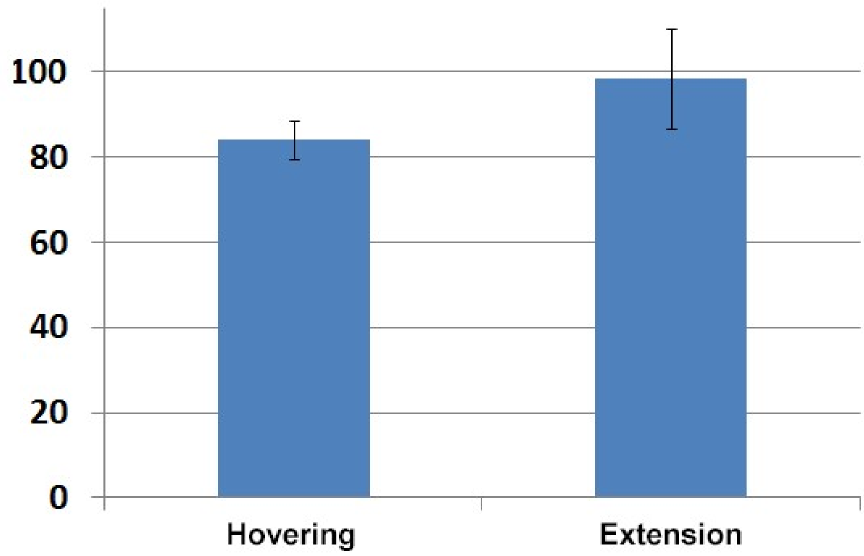


Figure 47: Results of task completion time (in seconds, error bar represent std. dev.).

other three (30%) answered the extension interface. This result reflected the results of the usability rating questions, while Chi-square test did not show significant difference from a random choice ($X^2(1) = 0.9$, $p = .343$).

Questions	Median [Q1 - Q3]		Wilcoxson Signed Rank Test
	Hovering interface	Extension interface	
Performed Well	5.5 [5 - 7]	5 [5 - 5.25]	Z = -2.121, p = .034*
Easy to Learn	6 [6 - 7]	5.5 [4 - 6.25]	Z = -1.807, p = .071
Easy to Use	6 [4.75 - 7]	5 [4 - 6]	Z = -1.381, p = .167
Intuitive	6 [6 - 7]	6 [4 - 6.25]	Z = -1.983, p = .047*
Natural	6 [4.75 - 6.25]	5 [4 - 5]	Z = -2.308, p = .021*
Effective	5.5 [4 - 6.25]	5 [4 - 6]	Z = -0.604, p = .546
Efficient	5 [4 - 6]	4 [4 - 5]	Z = -0.787, p = .431
Mental Stress	2 [1 - 3]	3 [2 - 5]	Z = -2.414, p= .016*
Physical Stress	5 [2.75 - 5.25]	4 [3 - 5]	Z = -1.029, p = .303

Table 7: Results of usability questions answered in 7-point Likert scale rating. (1: Not at all 7: Extremely, *: statistically significant difference)

When asked about the advantages of the hovering interface, the most common answer from participants was that it felt more natural and intuitive (5), followed by feeling more in control so that they could exactly know where the cursor will fall on the screen (4). Other answers included faster task performance and the tracking of the hand being more reliable.

When asked about what could be improved in the hovering interface, four participants replied nothing, while three of the participants mentioned that the tracking accuracy needed to be improved. Two participants mentioned being limited with the range where the user can reach, and one mentioned getting tired from stretching his arms to reach far away targets.

When asked what the advantages of the extension interface was, most of the participants (8, 80%) answered that it required less physical effort and movement. Two of the participants mentioned the method was fun to use.

For the question asking what needs to be improved with the extension interface, half of the participants (5) mentioned that the extension factor needed to be adjusted; most asking the extension to be made more subtle. As well as optimising the extension factor, two of the participants mentioned that the tracking accuracy needed to be improved.

Note that the study setup was in favour of the plain hovering method as the participants were standing at a distance where they can reach the far corners on the screen when stretching their arms. The results might be different with larger screens, in cases where the targets are out of reach of the participant.

8.4 SUMMARY OF THE MAPPING STUDY

We conducted a user experiment to investigate a non-linear extension method for controlling a gesture interface cursor.

Our results showed that the proposed method has above average usability, although improvements were possible in terms of adjusting the parameters that control the magnitude of the extension. One way to do this is by increasing the size of the region where the direct mapping is applied, and reducing the factor of extension to make the method to be more usable. The extension interface also showed promise in situations where the target is out of the user's reach, and where the hovering interface would not be suitable [Carroll and Thomas, 1988].

9 THE TARGET STUDY

The results from our mapping study provided useful feedback on how to improve the interface, notably that the magnitude of the extension in the extension interface was excessive, resulting in overshooting of the target and requiring participants to be more careful.

From that, we reason that by reducing the magnitude of cursor extension we can make the extension interface more usable. We also note that the experimental setup in the mapping study was leaning in favour of the hovering interface, as all of the targets were within the arms' reach of the participants.

We argue that there will be certain scenarios where the targets on the screen is out of arm's reach for the user. This can occur is a a user is situated a distance away from the screen, such that their arm reach cannot cover the entire screen, or if there are numerous objects on the interface, requiring placements away from the centre of a screen, where a users can easily reach. In such cases, the extension interface might prove more suitable than the hovering interface.

9.1 A DESIGN ITERATION

Taking our user feedback into consideration, and in the spirit of the iterative design process, we implement a design iteration to our prototype, and re-evaluate its usability after these improvements.

For the development of our prototype for this target study we recruited 5 users to test our modifications to the extension (non-linear mapping) interface, which was described in Chapter 7. Users were asked to interact with our prototype and comment on its usability as we adjusted the scaling factor, until we discovered one that they felt was the most comfortable or optimal.

By implementing the new mapping method developed in response to user feedback, we reason that the interface will feel more intuitive than our previous version.

9.1.1 A Pointing Interface

In this target study, we also evaluate a third method of interaction- the use of a pointing vector to control a cursor's position on the interface.

There are various ways to implement a pointing vector. To find out the most intuitive method, we conducted a quick user survey to explore the ideal reference points for a pointing vector. We interviewed on 8 participants (all male), asking them to point at various objects on a wall. The majority (5 out of 8) participants consistently pointed with a straight arm. The rest pointed with a bent elbow, or did a mixture of two. Of the participants, all 8 pointed by aligning their eye and hand to the pointing target. This was consistent with the findings by [Cheng and Takatsuka, 2009a] who found that the full arm stretch was the most common pointing strategy, while the most accurate strategy was when users lined up the target with their eye and fingertip. The basis of the pointing vector of our pointing interface is shown in Figure 48.

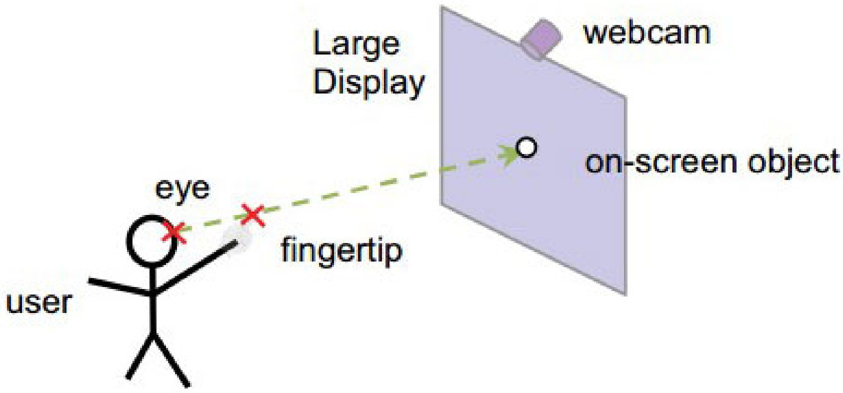


Figure 48: Reference points for calculating a pointing vector used to implement our pointing interface. From [Cheng and Takatsuka, 2009b].

9.2 THE SETUP

For the target study we used a rear screen projection setup, using translucent screen 2.4 meters wide and 1.8 meters high, mounted 0.60 meters above the ground. A projector is setup 2 meters behind the screen, running our custom software built for this study. A Kinect depth sensor is positioned 0.50 meters in front of the screen and mounted on a pole 1.2 meters from the ground. The setup is shown in Figure 49.

Figure 49: The setup for our target study. A screen displays our target boxes in a 6x4 configuration with a Kinect sensor mounted in front.

We utilise the Kinect's RGB camera to provide a live video feed for the mirror visualisation, as well as for its motion capturing ability. Our software, build in Unity, features a 6x4 grid of target boxes, coloured in yellow. The user's cursor is represented as a red circle. A random box is highlighted in blue to signify an active target. When the cursor stays on a target for 3 seconds the target is selected, disappears and another random target appears. A growing red box inside the target provides feedback on the time needed before it is selected. A timing system records the time taken to select the targets. There are 20 targets to select. We categorise the targets into 3 conditions- close targets, far targets, and a mixture of both. Distances were categorised relative to a user positioned in the centre of the screen, i.e. the Physical Interaction Zone (described in Chapter 3.3). See Figure 50 for a screenshot showing the targets to the user in the various conditions to be tested.

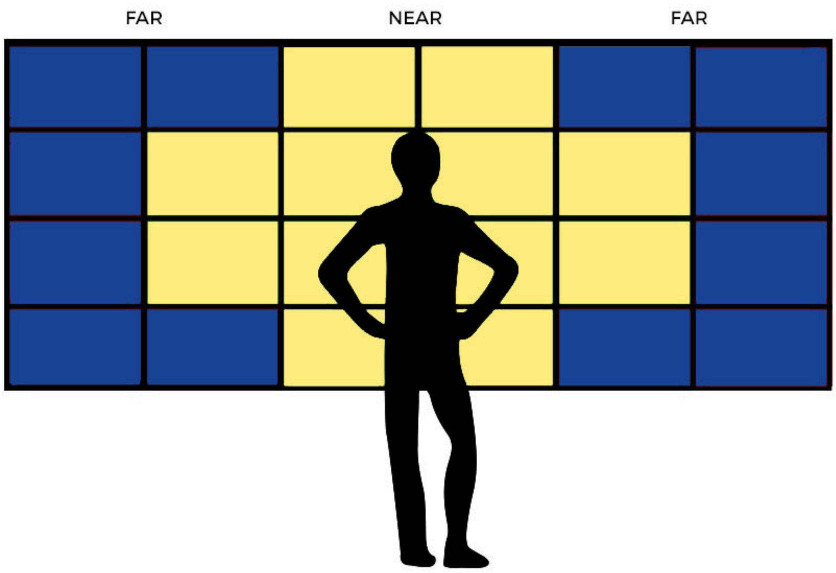


Figure 50: In the target study we tested the performance of each interface when selecting near, far and mixed (both near and far) proximity targets.

The software is also able to recognise 3 interaction methods- our hovering interface, the (modified) extension interface, and our new pointing interface. Both hands are able to interact, and in this study the user can choose to use one or both hands at once. A diagram showing difference between the interfaces is shown in Figure 51.

9.3 THE TASK

For this target study, the aim is for participant to select the blue target boxes. The time taken to select 20 targets is recorded. There are 3 sets of 24 boxes, making up the three different conditions. They consist of either a:

- random selection of near target,
- random selection of far targets,

- random selection of any (mixed) targets.

Note that only one target appears at a time. Once all 20 targets are selected, a task completion message is displayed and the task completion time is recorded. Participants repeat the task for each of the 3 conditions, in a randomised order. After each task, the participant completes the task questionnaire.

9.4 METRICS RECORDED

Questionnaires contain a range of open and closed questions to obtain demographic information on the participants and to get feedback on the experiment. The questionnaires answers are based on Likert scales ratings to evaluate the usability of the conditions and interactions used in the experiments. Task completion time is also recorded for each task, measured in seconds.

Our questionnaire adapted questions from the both the System Usability Scale [Usability.gov, 2015] and Nielsen's Attribute of Usability [Perlman, 2015] to provided a rounded approach to evaluating not only the system usability but aspects of the user interaction as well.

9.5 THE PROCEDURE

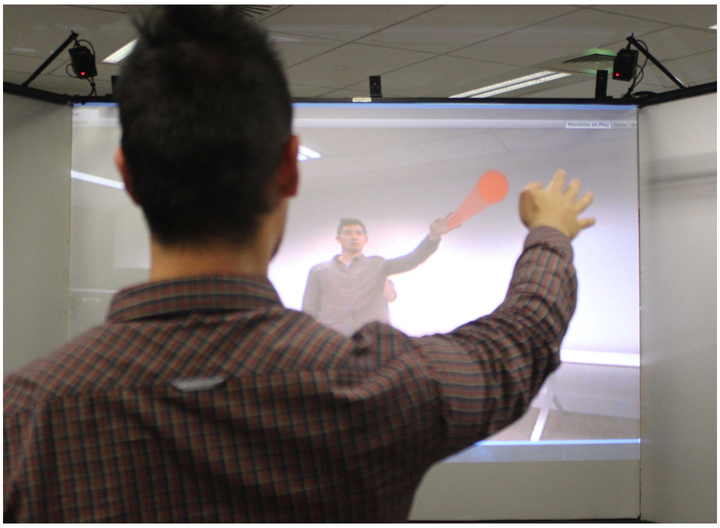
The experiment follows a similar procedure to the previous studies. Participant are first introduced to the objective- to explore the usability of the different mapping interfaces. They are then given a consent form to read and sign and a pre-experiment questionnaire to answer.

Participants are then asked perform the task, repeated for the three interfaces:

- direct mapping (hovering)
- extension (non-linear mapping)
- pointing (3D vector based)

Each participant is initially given a chance to interact freely with each interface to get used to system. The order in which the interface is tried is randomised. Once all interfaces have been tested, the participant answers the post-experimental questionnaire and gets a debriefing session to clarify any issues in the questionnaire responses.

(a) Hovering interface



(b) Extension interface

(c) Pointing interface

Figure 51: The different mappings of the interfaces, when the user's hand position in the mirror image is fixed.

10 RESULTS FROM THE TARGET STUDY

We analyse our data with a mixture of quantitative and qualitative feedback. As we are testing multiple factors (the interfaces and target proximity), for the task completion time we run a 3x3 two way repeated measures ANOVA with both factors as within group. For questionnaire results we run a Friedman's test with post hoc Wilcoxon Signed Ranked test with Bonferroni correction. The alpha level, is set to 0.05 for all of these test, unless otherwise stated.

10.1 PARTICIPANT ANALYSIS

We recruited 20 participants for this study, 11 male and 9 female, with an age range from 18 to 33 years old (*M*=23.4). Participants had various cultural backgrounds, with twelve participants identifying English as their first language and the rest ranging from Hindi, Marathi, Portuguese, Mandarin, Cantonese, Korean and German. All were proficient in English. Eight participants had never used a free hand gesture based interface before, while twelve replied that they have used it a couple of times a year.

10.2 TASK COMPLETION TIME ANALYSIS

The results from the task completion times are shown in Figure 52. Average task completion time for near targets was fastest for the hovering interface at 84.5 seconds (*SD*= 4.5), followed by the extension interface at 89.32 seconds (*SD*=10.9), with the pointing interface taking the longest at 126.1 seconds (*SD*=19.3).

For far targets, average completion time was fastest with the extension interface at 94.0 seconds (*SD*=7.5), followed closely by the hovering interface at 95.4 seconds (*SD*=15.15), with the pointing interface the longest at 125.2 seconds (*SD*=18.2).

With mixed proximity targets, the hovering interface was the fastest with an average of 92.52 seconds (*SD*= 9.22), followed by the extension interface at 92.13 seconds (*SD*=5.2) and lastly by the pointing interface at 121.8 seconds (*SD*= 14.6).

There was a significant main effect of the interface conditions, with the test statistic $F(1.19, 21.47) = 81.33$, $p < 0.001$. That is, ignoring other variables, task performance times were difference between the hovering, pointing and extension interfaces. Post-hoc pairwise comparison (with Bonferroni correction) showed a significant difference between the pointing and hovering interfaces ($p < 0.001$), and between the pointing and extension interface ($p < 0.001$). There was no significant difference between the hovering and extension interface ($p = 1.00$).

There was no significant main effect of proximity, $F(2, 36) = 2.54$, $p < 0.93$. That is, ignoring other variables, performance times were similar for near, mixed and far targets. Performance times of the different interfaces did not

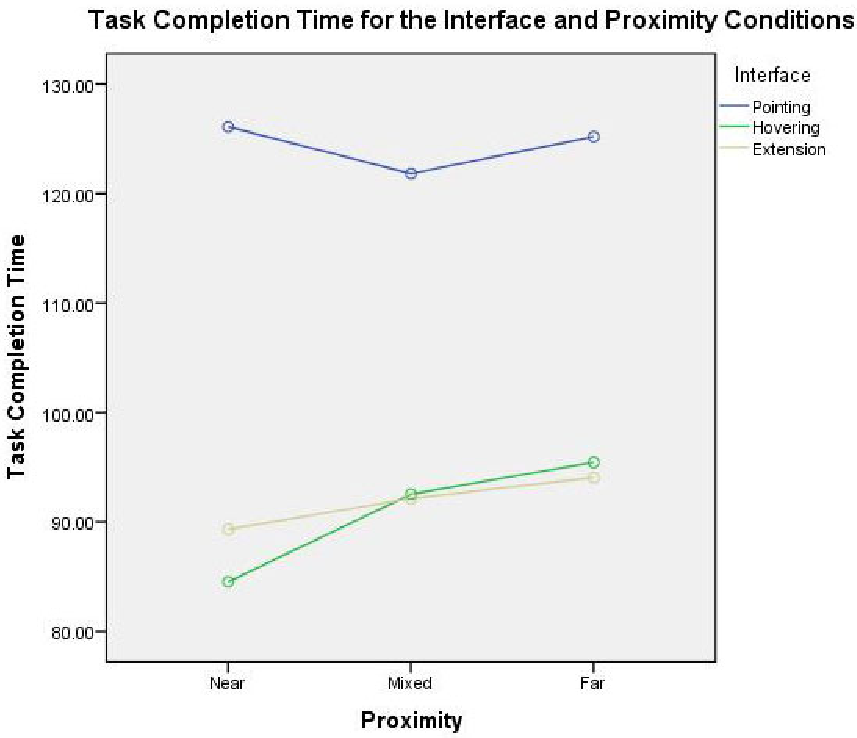


Figure 52: Average Task Completion Time for different target proximity (near, mixed, far), between the difference interfaces (pointing, hovering, and extension).

interact with the target proximity, F (4, 72) = 2.35, p< 0.06. This indicates that there is no effect on performance time from target distances between the pointing, hover and extension interfaces.

In summary, the pointing interface had a significantly longer task completion time compared to the hovering and extension interfaces. Task completion time between the hovering and extension interfaces showed no significant differences. Target proximity did not significantly affect the task completion time for all interfaces.

10.3 QUESTIONNAIRE ANALYSIS

We break down the questionnaire analysis into two main parts. In the first, we summarise the usability questionnaires regarding the task performance between the interfaces. In the second part we report on the overall ranking of the interfaces.

10.3.1 Usability

In this part we asked twelve questions on the usability of the interfaces. Questions had an answer consisting of a 7 point Likert scale, starting at 1 = Not at all, to 4 = Moderately, to 7 = Extremely. A detailed breakdown of each question is reported as follows, with a summary is shown in Table 8.

"I was able to perform the gesture well"

Statistically significant result, $X^2(20) = 11.36$, $p = 0.003$. Post hoc Wilcoxon signed-rank test with a Bonferroni correction (α= 0.017) showed significant differences between the hovering vs pointing interface (Z =-2.66, p=0.008) and extension vs pointing interface (Z=-2.57, p=0.010). No significance difference was found between the hovering versus extension interfaces (Z=-1.18, p= 0.24).

In summary, users felt they could perform gestures well in the hovering and extension interfaces (similarly), than the pointing interface.

"The gesture was easy to perform"

Statistically significant result, $X^2(20) = 12.47$, p =0.002. Post hoc Wilcoxon signed-rank test with a Bonferroni correction (α= 0.017) showed significant differences between the hovering vs pointing interface (Z=-3.151, p=0.002) and extension vs pointing interface (Z=-2.727, p=0.006). No significances difference was found between the hovering versus extension interface (Z=-0.97, p=.33).

In summary, users felt the gestures were easy to perform in the hovering and extension interfaces (similarly), than the pointing interface.

"The gesture is intuitive"

Statistically significant result, $X^2(20) = 9.73$, p=0.008. Post hoc Wilcoxon signed-rank test with a Bonferroni correction (α= 0.017) showed significant differences between the hovering vs pointing interface (Z=-2.66, p=0.008). No significances difference was found between the hovering vs and extension interface (Z=-1.38, p=.167), or the extension vs pointing interface (Z=-1.81, p=0.070).

In summary, users felt the gestures in the hovering interface was the more intuitive than those in the pointing interface. The gestures in the extension interface did not feel significantly different from either.

"The gesture feels natural"

Statistically significant result, $X^2(20) = 8.69$, p=0.013. Post hoc Wilcoxon signed-rank test with a Bonferroni correction (α= 0.017) showed no significant differences between the extension vs pointing interface (Z=-1.13, p=0.26), the hovering versus pointing interface (Z=-2.21, p=0.027), or the hovering versus extension interface (Z=-2.15, p=0.03).

In summary, users felt the gestures were natural in all three interfaces, similarly.

"The gesture is fun to perform"

No statistically significant result, $X^2(20) = 1.81$, p=0.41.

"The gesture feels accurate"

Statistically significant result, $X^2(20) = 16.33$, p=< 0.001. Post hoc Wilcoxon signed-rank test with a Bonferroni correction (α= 0.017) showed significant difference between the hovering vs pointing interface (Z=-3.369, p= 0.001). No significant differences found between hovering versus extension interface or the extension versus pointing interfaces.

In summary, users felt the gestures were more accurate in the hovering interface than the pointing interface. The extension interface did not feel significantly different from either in respect to gesture accuracy.

"The gesture is efficient to do the given task"

No statistically significant result, $X^2(20) = 4.03$, p=0.13.

"The gesture is mentally stressful"

Statistically significant result, $X^2(20) = 12.28$, p=0.002. Post hoc Wilcoxon signed-rank test with a Bonferroni correction (α= 0.017) showed significant difference between the hovering vs pointing interface (Z=-3.09, p=0.002). No significance difference found between hovering versus extension interface, or the extension versus pointing interfaces.

In summary, users felt the gestures were more mentally stressful in the pointing interface than the hovering interface. The extension interface did not feel significantly different from either.

"The gesture is physically stressful"

No statistically significant result, $X^2(20) = 4.84$, p=0.089.

"The gesture is easily recognised by a person"

No statistically significant result, $X^2(20) = 4.79$, p=0.09.

"The gesture is easily recognised by the computer"

Statistically significant result, $X^2(20) = 11.33$, p=0.003. Post hoc Wilcoxon signed-rank test with a Bonferroni correction (α= 0.017) showed a significant difference between the hovering vs pointing interface (Z=-2.97, p=0.003). No significance difference found between the hovering versus extension interface, or the extension versus pointing interface.

In summary, users felt the gestures were more easily recognised by the computer in the hovering interface than the pointing interface. The extension interface did not feel significantly different from either.

10.3.2 Interface Rankings

Here we report the post experiment questionnaire results where participants rank the overall usability of the three interfaces- pointing, extension and hovering.

Ease of Use

When ranking which interface was the easiest use, there was a statistically significant difference between the interfaces, $X^2(20) = 19.90$, $p< 0.001$. Post hoc analysis with the Wilcoxon signed-rank test was conducted with a Bonferroni correction applied, resulting in a significance level set a α= 0.017. There were no significant difference in rankings based on ease of use between the extension versus pointing interface (Z=-1.93, p=0.05), however there was a significant result in the ease of use between the hovering versus pointing interface (Z=4.05, p=< 0.001) and the hovering versus extension interface (Z=-2.71, p=0.007).

Question	Median [Q1 - Q3]		
	Hovering	Extension	Pointing
"I was able to perform the gesture well"*	6 [5.25 - 6][a]	5 [5 - 6][b]	5 [4 - 5.75][a,b]
"The gesture was easy to perform"*	6 [5 - 7][a]	6 [5 - 7][b]	5 [3.25 - 5][a,b]
"The gesture is intuitive"*	6 [5 - 6][a]	5.5 [4 - 6]	4.5 [3 - 6][a]
"The gesture feels natural"*	6 [5 - 6.75]	5 [4 - 6]	5 [3.25 - 6]
"The gesture is fun to perform"	5 [4 - 6]	5.5 [4.25 - 6]	4 [4 - 5.75]
"The gesture feels accurate"*	5 [5 - 6][a]	4 [4 - 5.75]	4 [3 - 5][a]
"The gesture is efficient to do the given task"	5 [4.25 - 6]	5 [4 - 6]	5 [4 - 5]
"The gesture is mentally stressful"*	1 [1 - 2][a]	2 [1.25 - 3]	3 [2 - 4][a]
"The gesture is physically stressful"	2 [1 - 4]	2 [1 - 3]	3 [1.25 - 4]
"The gesture is easily recognized by a person"	6 [5 - 6.75]	5 [4 - 6]	5 [4 - 6]
"The gesture is easily recognized by the computer"*	6 [5 - 6.75][a]	5.5 [4 - 6]	5 [3.25 - 5][a]

Table 8: Results of usability questions answered in point Likert scale rating. (1: Not at all 7: Extremely). Significant results are marked with an asterisks (*) and the superscript (e.g. a, b) denotes the pairs that show significant difference.

In summary, users ranked the hovering interface as the easiest to use, with the pointing interface as the hardest to use. A summary is shown in Figure 53.

Most Intuitive

When ranking which interface was the most intuitive to use, there was a statistically significant difference between the interfaces, $X^2(20) = 13.30$, p=0.001. Post hoc analysis with Wilcoxon signed-rank test was conducted with a Bonferroni correction applied, resulting in a significance level set at α= 0.017. There was no significant difference between the rankings based on intuitiveness of the extension versus hovering interface (Z=-1.56, p=0.12) or the extension versus pointing interfaces (Z=-2.17, p=0.03). The pointing versus hovering interface showed a statistically significant result, with the hovering being more intuitive (Z=-3.35, p=0.001).

In summary, users ranked the hovering interface as the most intuitive, with the pointing interface as the least intuitive. A summary is shown in Figure 54.

Fun to Use

When ranking which interface was the most fun to use, there was a statistically significant difference between the interfaces, $X^2(20)$= 34.30, $p < 0.001$. Post hoc analysis with Wilcoxon signed-rank test was conducted with a Bonferroni correction applied, resulting in a significance level set at α= 0.017. There were significant differences between the rankings based on fun in the pointing versus hovering interface (Z=-4.23, $p < 0.001$), the extension versus hovering interface (Z=-3.545, $p < 0.001$) and the extension versus pointing interface (Z=-3.91, $p < 0.001$).

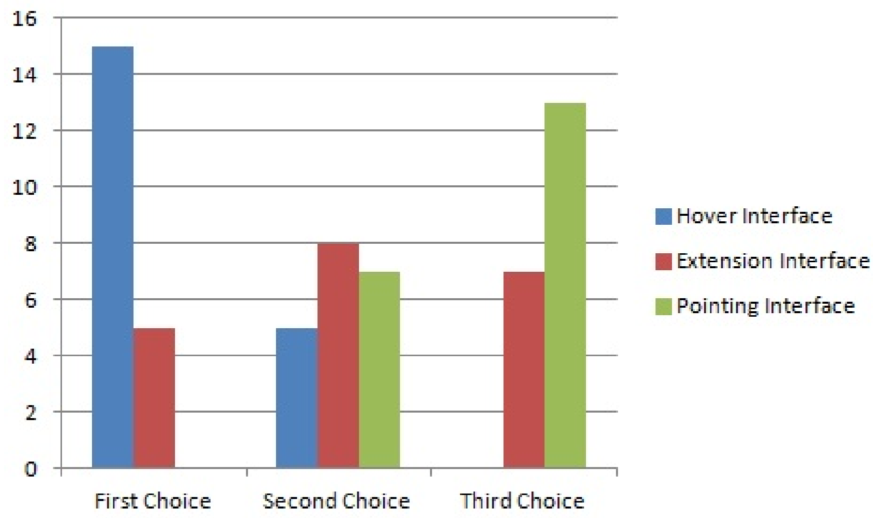


Figure 53: Interface rankings: Participants rank the three interfaces into which was the easiest to use.

In summary, users ranked the hovering interface as the most fun to use and the pointing interface as the least fun, with the extension interface in between. A summary is shown in Figure 55.

Most accurate

When ranking which interface was the most accurate to use, there was a statistically significant difference between the interfaces, $X^2(20) = 13.30$, p=0.001. Post hoc analysis with Wilcoxon signed-rank test was conducted with a Bonferroni correction applied, resulting in a significance level set at α= 0.017. There was a significant difference between in rankings based on accuracy in the hovering versus pointing interface (Z=-3.35, p=0.001). There was no statistical significance for extension versus hovering interface (Z=-1.56, p=0.12) or the extension versus pointing interface (Z=-2.17, p=0.03).

In summary, users ranked the hovering interface as the most accurate to use, with the pointing interface as the least accurate. The extension interface was perceived to be similar in accuracy to both. A summary is shown in Figure 56.

Most preferred

When ranking which interface was the most preferred, there was a statistically significant difference between the interfaces, $X^2(20)$ = 7.30, p=0.03. Post hoc analysis with Wilcoxon signed-rank test was conducted with a Bonferroni correction applied, resulting in a significance level set at α= 0.017. There were significant differences between the preference for the extension versus the pointing interface (Z=-2.656, p=0.008). There was no statistical significance for the pointing versus hovering interface (Z=-0.73, p=0.47) or the extension versus hovering interface (Z=-1.95, p=0.05).

In summary, users preferred the extension interface over the pointing interface, although they had no significant preference between the hovering vs extension or hovering vs pointing interfaces. A summary is shown in Figure 57.

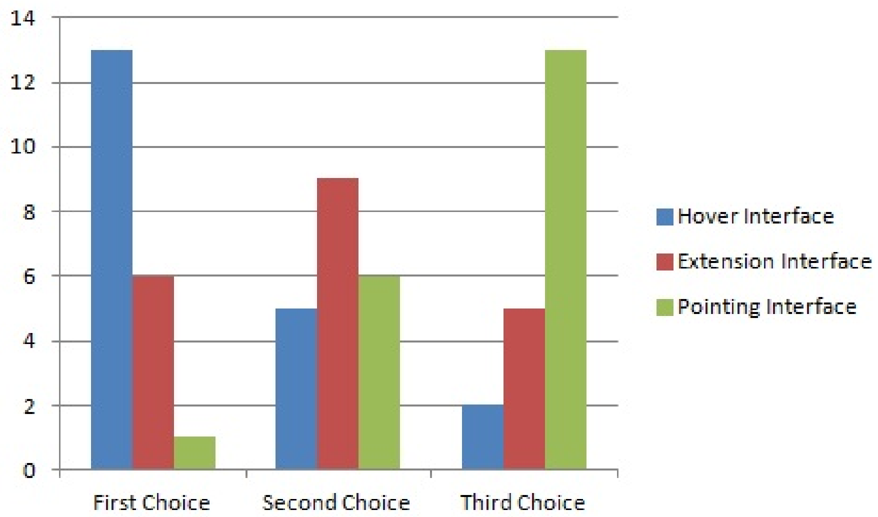


Figure 54: Interface rankings: Participants rank the three interfaces into which was the most intuitive to use.

10.3.3 Interface Comments

Comments about the hovering interface included remarks about its ease of use, with one participant saying that it is "intuitive and easy to pick up," and another commenting that it required "little guessing or learning where the cursor is." A common downside of the hovering interface is that "you have to move a lot," leading to some preferring the extension interface, saying "I don't have to move too much." Interestingly, one participant commented that the extension interface was "cool but I like moving around the room!" and favoured the hovering interface instead. While all the interfaces had similar fun rankings, the extension interface attracted more comments, being compared to "having superpowers like Spiderman" or "having a yo-yo in your hand." When asked about the accuracy of the interface, participant found the hovering interface "very accurate" compared to the other methods. Interestingly, one participant concluded that the pointing method was inaccurate as it was "too sensitive" and followed that the hovering interface must have been the "least sensitive as it was the most accurate."

10.4 SUMMARY OF THE TARGET STUDY

In summary, our target study explored the use of three interfaces, each using different mapping methods for controlling a cursor on a gesture interface. We also tested the effect of selecting targets at different proximity to the user.

The task completion times were similar for the hovering and extension interfaces, with average task completion time for both interfaces being faster than the pointing interface. The extension interface had the fastest task completion time for selecting far targets, although target proximity was not found to be a significant factor in this study.

The hovering interface was found to be easier, more intuitive, and more accurate than the pointing interface, while the extension interface performed similar to the hovering interface in these aspects, with no statistically signif-

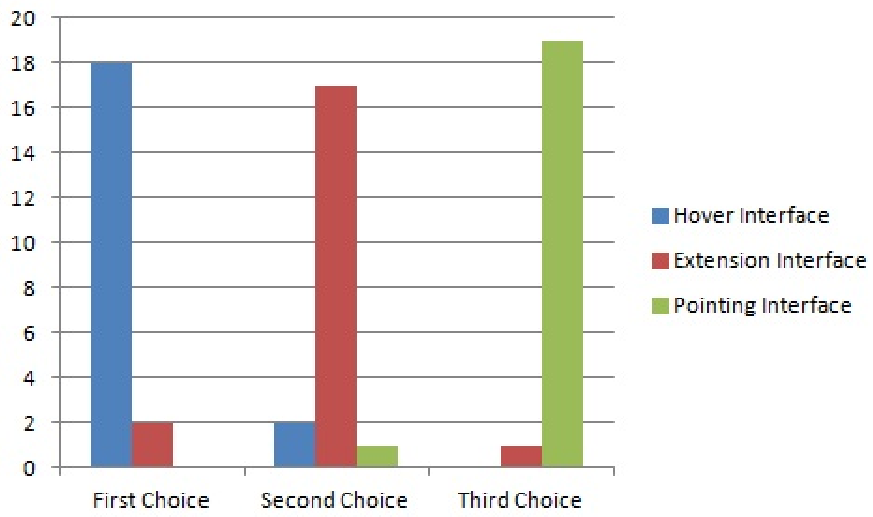


Figure 55: Interface rankings: Participants rank the three interfaces into which was the most fun to use.

icant difference. This suggests that the modified mapping method in our extension method managed to increase the usability of our extension interface, providing a user experience to rival the hovering interface.

Pointing was the least preferred interface, although there was no significant difference between the hovering and pointing interface.

The accuracy, physical stress and mental stress of using the hovering and extension interfaces showed no significant difference, although both methods outperformed the pointing interface. When asked about how natural the different interfaces felt, all were perceived to be of similar, and all were ranked above average.

10.5 DISCUSSION ON THE TARGET STUDY

The pointing interface took the longest time to complete the task. Poor performance of the pointing interface may have been caused by implementation issues, although questionnaire evaluation of the interface's accuracy and perceived ability to be recognised by the computer showed that it performed similarly to the extension interface, making this issue unlikely.

The hovering interface had similar task completion times to the extension interface, also showing no significant difference in the physical or mental stress, intuitiveness, accuracy or preference between the two interfaces. Difference between the two interfaces were identified when rankings for ease of use and perceived fun, where the hovering interface ranked higher than the extension interface.

It is interesting to note that despite no difference in the ease of use between the extension and pointing interfaces, a difference was shown in their ranking for fun. The correlation of ease of use and fun is explained by [Carroll and Thomas, 1988], who suggests that the "ease" is not necessarily an indicator of "fun" and that fun can arise from things that have moderate complexity, which promotes user learning and engagement.

When evaluating accuracy, users felt their gestures were less accurate when using the pointing interface, compared to the hovering interface. The

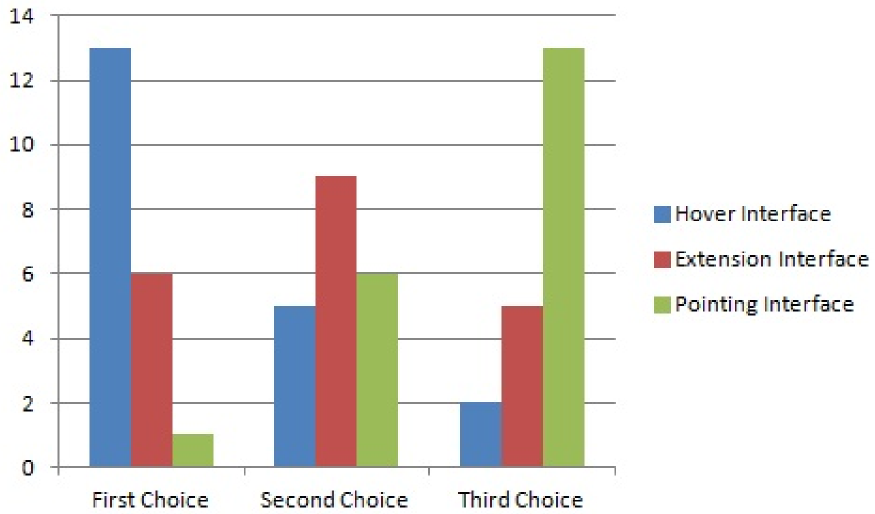


Figure 56: Interface rankings: Participants rank the three interfaces into which was the most accurate to use.

perception of accuracy may have been affected by an increased demand in concentration needed for controlling the cursor in the pointing interface, as users also felt the pointing interface was mentally stressful compared to the hovering interface, indicating a psychological component in perceiving accuracy. Some participants also felt that the pointing interface was too sensitive, and questioned the experimenter if the cursor was detecting their shaky hands. This could be fixed by decreasing the sensitivity of the pointing vector (e.g. incorporating some sort of low pass filter algorithm).

Overall ranking of preference placed the pointing interface last, with no significant difference in ranking between the hovering and extension interface. In our background review in Chapter 2.6, we found that interactive gesture systems are commonly use in a public setting- as such, there is limited time for learning and interacting. Therefore when implementing our prototype we aimed for a task completion time of one to two minutes. While short enough to show a difference in the pointing interface, the lack of distinction between the results of task completion time between hovering and extension interface might be clearer if we increase the interaction time (i.e. longer task). The lack of any significant difference between selecting targets of varying proximity on the screen could also be attributed to the short task interaction time, where total energy expenditure is minimal. In their questionnaire response, all interfaces showed no significant difference in physical stress, despite their different physical requirements. Possible factors that are able to reduce the perception of fatigue when interacting with a gesture interfaces include:

- having a short interaction time (approx two minutes)
- ability to use both arms
- a 'gaming' attitude to the interaction
- the novelty of a gesture system

While the extension interface was designed to be the less physically demanding interaction method, this study was not set up to specifically test

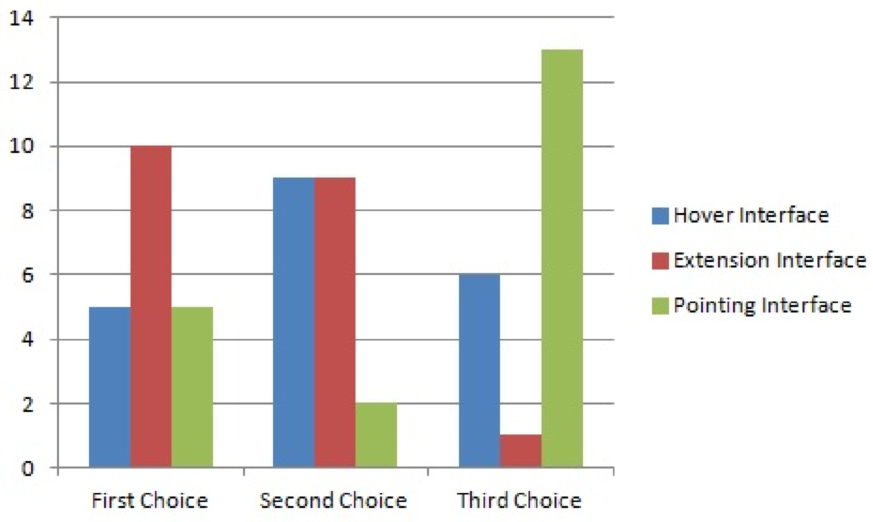


Figure 57: Interface rankings: Participants rank the three interfaces into which was the most preferred to use.

this. Further research could investigate it's effectiveness by implementing a longer task completion time, restricting users to one hand for interaction, or even recruiting users who are used to interacting with gesture interfaces. It is also interesting to note that [Rico and Brewster, 2010] found energy efficiency to be of low priority to users when determining the usability of a gesture.

11 OVERALL DISCUSSION AND CONCLUSIONS

In this chapter we discuss the overall findings from our three experiments, outlining a set of design guidelines for developing gestural interfaces before suggesting areas for future research.

11.1 RESEARCH SUMMARY

We started our research with the aim of developing gestures for interactive large screen displays. In our background review (Chapter 2) we find the growing use of gesture technology for applications in the retail, health and fitness, and in the public domain. We then summarise the main factors to consider when developing a gesture interface in Chapter 3, before developing our prototypes with the tools described in 4.

We conducted a guessability study in Chapter 5 to discover user-defined gestures for common tasks encountered when interacting with a gesture interface. We discovered that participants mainly used deictic and metaphoric gestures, and for certain interaction tasks there was generally a consensus on what the most natural gesture to use was. For tasks where agreement is lower between participants, developing a system able to recognise multiple gestures for one action should increase the user experience. An alternative would be to use visual cues, although our literature review suggests that it might not be as effective.

The use of hovering gestures (with both a linear and non-linear mapping of the user's hand) was preferred over pointing gestures when interacting with our gesture interface. This could have been due to the close proximity of the targets, although in our target study, we found that turning off the mirror visualisation led to all participants favouring pointing gestures. The number of participants used for the sub-study was too small to produce any significant conclusions and would be a topic for future research. Another factor that might have led to this was the augmented mirror visualisation present in our interface, which used a streaming video camera as the background. This mirror visualisation provided a dynamic representation of the user that may have given our interface certain affordance properties. This could be due to the presence of a dynamic mirror representation hinting to users that their body movement is being recognised by the system. One could also argue that the augmented reality visualisation provided another dimension to the setup, the concept of a real world and virtual world in the mirror image, connected through a large screen display. As such their gesture may reflect a subconscious response to controlling their mirror 'avatar' in a way as to interact with the virtual object like a puppeteer would to his or her characters. This concept of two different worlds connected by an interface has been used in marketing campaigns, such as the Coca-Cola Small World Machines [Coca-Cola, 2013]. The company installed touch interfaces in two locations, one in India and one in Pakistan. Both countries had a tense political relationship at the time, and the machines provided a way to connect and engage users between the two location. It is shown in Figure 58.

Figure 58: In Coca-Cola's Small World Machines, an interactive display is installed in India (left) and Pakistan (right). The interface (centre) has a live video stream and incorporates co-operative games to engage users in both location. From [Coca-Cola, 2013]

The ways users scroll on a gesture interface also gave rise to an interesting observation- how do you gesture to start and stop scrolling? There was a lack of consensus among participants. Some tried a pushing motion for initialisation, some waited, and some even tried hiding their hands behind their back. This has been a long standing problem for gesture interfaces, one that has been called the Midas touch effect [Kortum, 2008]. In Greek mythology, King Midas wished for the ability to turn anything he touched into gold. This seemed a blessing until he realised that everything he touch- his food, drink, and even his daughter, turned to gold. With the constant tracking of a user movement in gesture interfaces, confusion can arise as to when a gesture should be considered started or stopped, and a user centred approach to this problem is warranted.

Through the pointing task in our study we discovered that the positioning of a cursor is an important element of a gesture interface, as it is on any other computer interface. The use of hovering was not effective for some users who did not position their hand in alignment with the objects on the screen, suggesting the need for more convenient way to reach targets. From this we introduced an extension interface using a non-linear mapping of the cursor, and tested our prototype in our mapping study (Chapter 7).

Our mapping study showed that while our extension interface had an above average usability, it still needed some modifications to improve the user experience. After making changes to our prototype, we evaluate its performance in our target study (Chapter 9).

In our target study study we introduced an additional interaction method, the pointing interface, which positioned the cursor via a pointing vector. Figure 59 shows an overview of our three interfaces. While pointing may seem to be the most natural thing to do when selecting objects in the real world, in human-computer interaction, the use of the word "natural" can be thought of as a relative term, specific to an application or task. Using a computer itself may not be considered natural; and while using gestures can be considered more natural than using a keyboard, for the task of writing a long text document, the keyboard prevails. The hovering and extension interfaces outperformed the pointing interfaces. This could be due to the higher cognitive threshold needed to mentally calculate the cursor position in the pointing interface. When using the hovering and extension interface, participants could adjust the position the cursor using a combination of:

- their own body orientation

- the cursor position on the screen
- position of their avatar in the augmented mirror representation.

In the pointing interface however, this last factor, the augmented mirror representation, did not provide relevant feedback into position of the cursor (as the pointing vector is calculated only from their body position). As such, participants had less sensory data to mentally align the cursor, which led to slower performance times.

[Tractinsky, 1997] tells us that the usability of a system is not purely derived from their aesthetic value, but from the whole experience, i.e. the cognitive stimulation of the user- highlighting one of the benefits of developing with the user in mind. By following this in our research, we have demonstrated that a user centred approach to developing gesture interfaces can lead to a system with superior usability and user experience, and the implementation of an extension interface has shown potential as a standard interaction method for gesture interfaces on large screen displays.

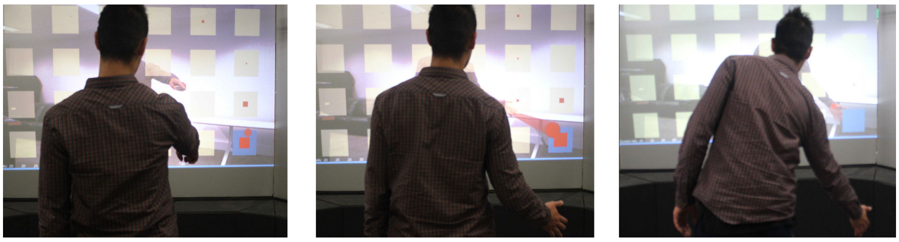


Figure 59: The different movements needed when selecting the same target. The user has to move more in the hovering interface (right) compared to the extension (centre) and pointing (left) interfaces.

11.2 DESIGN GUIDELINES

In this research we followed a user-centred approach for the development of gesture interactions with a large screen display. By evaluating the usability of gestures we identified factors that can increase a user's experience and will be of benefit to future developers of gesture interfaces. The main points are:

- **accept multiple hand poses**- e.g. an open hand, closed hand and various finger poses should be capable of completing the same interaction tasks.
- **accept multiple gestures**- some interaction tasks (like choosing between dichotomous options) can be achieved with various gestures. This is analogous to a desktop computer, where you can delete an object with either the DELETE key or drag and drop it into the recycling bin icon. Users are no homogeneous and therefore system should not necessarily be.
- **dynamic representations are engaging**- utilising a mirror interfaces provides a good method for conveying a system's interactivity and can improve gesture agreement between users.
- **short interaction time**- limiting continuous interaction times to approximately one to two minutes limits the impact of fatigue on users.

- **allow two handed interaction**- allows users to switch their hands if tired, and to interact with the system when only one hand is available.

11.3 FUTURE WORK

When the augmentation effect was removed, users switched to predominantly pointing gestures, as one would gestures to physical objects in the real world. Our sample size was too small to draw any significant conclusions from this no-mirror effect, but it suggests a direction for future research.

Other potential areas to explore would be the use of these user defined gestures in an in-the-wild study, where our gesture interfaces are used in a public setting. Gestures, by nature, are social constructs, and it will be interesting to explore how users interact with a gesture interface when there are multiple users or bystanders around.

BIBLIOGRAPHY

Accenture

2006 *New Accenture Technology Lands at O'Hare International Airport*, Web Page, https://newsroom.accenture.com/article_display.cfm?article_id=4341.

Ackad, Christopher, Judy Kay, and Martin Tomitsch

2014 "Towards Learnable Gestures for Exploring Hierarchical Information Spaces at a Large Public Display," in *CHI'14 Workshop on Gesture-based Interaction Design, Toronto, Canada*.

Ackad, Christopher, Rainer Wasinger, Richard Gluga, Judy Kay, and Martin Tomitsch

2013 "Measuring interactivity at an interactive public information display," in *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration*, ACM, pp. 329-332, ISBN: 1450325254.

Annett, Michelle and Walter F Bischof

2013 "Your left hand can do it too!: investigating intermanual, symmetric gesture transfer on touchscreens," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, pp. 1119-1128, ISBN: 1450318991.

ARDOORMoscow

2011 *Kinect Fitting Room for Topshop*, Web Page, https://www.youtube.com/watch?t=134&v=L_cYKFdP1_0.

Brignull, Harry and Yvonne Rogers

2003 "Enticing people to interact with large public displays in public spaces," in *Proceedings of INTERACT*, vol. 3, pp. 17-24.

Brooke, John

1996 "SUS-A quick and dirty usability scale," *Usability evaluation in industry*, vol. 189, 194, pp. 4-7.

Buerger, Neal

2011 "Types of public interactive display technologies and how to motivate users to interact," *Displays*.

Caron, Nathalie

2015 *Meet the rest of the cast for Fox's Minority Report TV series*, Web Page, http://www.blastr.com/2015-3-3/meet-rest-cast-foxs-minority-report-tv-series.

Carroll, John M and John C Thomas

1988 "Fun," *ACM SIGCHI Bulletin*, vol. 19, 3, pp. 21-24.

Cassell, Justine

1998 "A framework for gesture generation and interpretation," *Computer vision in human-machine interaction*, pp. 191-215.

Chang, Yao-Jen, Shu-Fang Chen, and Jun-Da Huang

2011 "A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities," *Research in developmental disabilities*, vol. 32, 6, pp. 2566-2570, ISSN: 0891-4222.

Cheng, Kelvin and Masahiro Takatsuka

2009a "Hand pointing accuracy for vision-based interactive systems," in *Human-Computer Interaction–INTERACT 2009*, Springer, pp. 13-16, ISBN: 3642036570.

2009b *Interaction Paradigms for Bare-Hand Interaction with Large Displays at a Distance*, Web Page, http://cdn.intechopen.com/pdfs-wm/8967.pdf.

Cisco

2011 *The Virtual Fashion Mirror is the future of shopping*, Web Page, https://www.youtube.com/watch?t=114&v=Q9iq20G3KLA.

Citi

2010 *Citi: Retail Banking Restyled*, Web Page, https://www.youtube.com/watch?v=9U3MZeku_eg.

Coca-Cola

2013 "Coca-Cola Small World Machines - Bringing India and Pakistan Together," https://www.youtube.com/watch?v=ts_4vOUDImE.

Control Group

2015 *WPC Interactive Installation*, Web Page, http://www.controlgroup.com/chevron.html.

Demodern

2014 *Affordance of public touchscreens*, Web Page, http://demodern.com/blog/public-touchscreens.

Grace, Kazjon, Rainer Wasinger, Christopher Ackad, Anthony Collins, Oliver Dawson, Richard Gluga, Judy Kay, and Martin Tomitsch

2013 "Conveying interactivity at an interactive public information display," in *Proceedings of the 2nd ACM International Symposium on Pervasive Displays*, ACM, pp. 19-24, ISBN: 1450320961.

Hespanhol, Luke, Martin Tomitsch, Kazjon Grace, Anthony Collins, and Judy Kay

2012 "Investigating intuitiveness and effectiveness of gestures for free spatial interaction with large displays," in *Proceedings of the 2012 International Symposium on Pervasive Displays*, ACM, p. 6, ISBN: 1450314147.

HITLab NZ

2011 , Web Page, http://www.hitlabnz.org/index.php/research/augmented-reality.

HotHardware

2015 , Web Page, http://hothardware.com/reviews/apple-ipad-or-a-netbooktablet-hybrid-which-makes-more-sense.

Human-Computer Interaction Institute

2011 , Web Page, http://hciresearch4.hcii.cs.cmu.edu/M-HCI/2011/GE-Cardiology/iterative-design.php.

IMGSRCinc

2013 *Genso Scope (element scope) installation*, Web Page, https://www.youtube.com/watch?v=8TbKGnp_0zo.

INC, NON-GRID

2014 *PUMA Store Harajuku Interactive Mirror*, Web Page, http://non-grid.com/puma-store-interactive-mirror/.

Jakobsen, Mikkel R, Yonas Sahlemariam Haile, Soren Knudsen, and Kasper Hornbæk

2013 "Information visualization and proxemics: design opportunities and empirical findings," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 19, 12, pp. 2386-2395, ISSN: 1077-2626.

Kortum, Philip

2008 *HCI beyond the GUI: Design for haptic, speech, olfactory, and other non-traditional interfaces*, Morgan Kaufmann, ISBN: 0080558348.

Kules, Bill, Hyunmo Kang, Catherine Plaisant, Anne Rose, and Ben Shneiderman

2003 "Immediate usability: Kiosk design principles from the CHI 2001 photo library."

Leap Motion

2015 "Leap Motion," https://www.leapmotion.com/.

Lee, Bongshin, Petra Isenberg, Nathalie Henry Riche, and Sheelagh Carpendale

2012 "Beyond mouse and keyboard: Expanding design considerations for information visualization interactions," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 18, 12, pp. 2689-2698, ISSN: 1077-2626.

McNeill, David

1992 *Hand and mind: What gestures reveal about thought*, University of Chicago Press, ISBN: 0226561321.

Mesterius

2013 *Teletoon Officially Announces Inspector Gadget's New CGI Series*, Web Page, http://nexttime-gadget.blogspot.co.nz/2013/06/teletoon-officially-announces-inspector.html.

Michelis, Daniel and Jörg Müller

2011 "The audience funnel: Observations of gesture based interaction with multiple large displays in a city center," *Intl. Journal of Human–Computer Interaction*, vol. 27, 6, pp. 562-579, ISSN: 1044-7318.

Microsoft

2013 *Kinect for Windows Retail Clothing Scenario*, Web Page, http://www.microsoft.com/en-us/showcase/details.aspx?uuid=960852fa-9bcd-4aad-b292-9b78cec28659.

2015a , Web Page, http://blogs.msdn.com/b/benriga/archive/2008/07/23/my-next-gadget-i-wish.aspx.

2015b *Kinect for Windows Human Interface Guidelines v1.8.0*, Web Page, https://msdn.microsoft.com/en-us/library/jj663791.aspx.

2015c *Kinect for Windows SDK*, Web Page, https://msdn.microsoft.com/en-us/library/hh855347.aspx.

Microsoft

2015d *Kinect for Windows Sensor Components and Specifications*, Web Page, https://msdn.microsoft.com/en-us/library/jj131033.aspx.

2015e *Skeletal Tracking*, Web Page, https://msdn.microsoft.com/en-us/library/hh973074.aspx.

2015f *Touchless Interaction in Medical Imaging*, Web Page, http://research.microsoft.com/en-us/projects/touchlessinteractionmedical/.

Müller, Jörg, Florian Alt, Daniel Michelis, and Albrecht Schmidt

2010 "Requirements and design space for interactive public displays," in *Proceedings of the international conference on Multimedia*, ACM, pp. 1285-1294, ISBN: 1605589330.

Müller, Jörg, Robert Walter, Gilles Bailly, Michael Nischt, and Florian Alt

2012 "Looking glass: a field study on noticing interactivity of a shop window," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, pp. 297-306, ISBN: 145031015X.

Müller, Jörg, Dennis Wilmsmann, Juliane Exeler, Markus Buzeck, Albrecht Schmidt, Tim Jay, and Antonio Krüger

2009 "Display blindness: The effect of expectations on attention towards digital signage," in *Pervasive Computing*, Springer, pp. 1-8, ISBN: 3642015158.

Nancel, Mathieu, Julie Wagner, Emmanuel Pietriga, Olivier Chapuis, and Wendy Mackay

2011 "Mid-air pan-and-zoom on wall-sized displays," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, pp. 177-186, ISBN: 1450302289.

Nielsen, Michael, Moritz Störring, Thomas B Moeslund, and Erik Granum

2003 "A procedure for developing intuitive and ergonomic gesture interfaces for man-machine interaction," in *Proceedings of the 5th International Gesture Workshop*, pp. 1-12.

Norman, Donald A

2013 *The design of everyday things: Revised and expanded edition*, Basic books, ISBN: 0465072992.

Norman, Donald and Jakob Nielsen

2010 *Gestural Interfaces: A Step Backwards In Usability*, Web Page, http://www.jnd.org/dn.mss/gestural_interfaces_a_step_backwards_in_usability_6.html.

NuFormer

2013 *Interactive window projection*, Web Page, (http://www.nuformer.com/portfolio/item/interactive-window-projection/).

Peltonen, Peter, Esko Kurvinen, Antti Salovaara, Giulio Jacucci, Tommi Ilmonen, John Evans, Antti Oulasvirta, and Petri Saarikko

2008 "It's Mine, Don't Touch!: interactions at a large multi-touch display in a city centre," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, pp. 1285-1294, ISBN: 1605580112.

Perlman, Gary

2015 , Web Page, http://garyperlman.com/quest/.

Piumsomboon, Thammathip, Adrian Clark, Mark Billinghurst, and Andy Cockburn

2013 "User-defined gestures for augmented reality," in *Human-Computer Interaction–INTERACT 2013*, Springer, pp. 282-299, ISBN: 3642404790.

Poupyrev, Ivan, Mark Billinghurst, Suzanne Weghorst, and Tadao Ichikawa

1996 "The go-go interaction technique: non-linear mapping for direct manipulation in VR," in *Proceedings of the 9th annual ACM symposium on User interface software and technology*, ACM, pp. 79-80, ISBN: 0897917987.

Rico, Julie and Stephen Brewster

2010 "Usable gestures for mobile interfaces: evaluating social acceptability," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, pp. 887-896, ISBN: 1605589292.

Rose, Brent

2012 *Nike+ Kinect Training Hands On: Some Serious Home Exercise*, Web Page, http://www.gizmodo.com.au/2012/10/nike-kinect-training-hands-on-some-serious-home-exercise/.

Ruiz, Jaime, Yang Li, and Edward Lank

2011 "User-defined motion gestures for mobile interaction," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, pp. 197-206, ISBN: 1450302289.

Rydén, Fredrik, Howard Jay Chizeck, Sina Nia Kosari, Hawkeye King, and Blake Hannaford

2011 "Using kinect and a haptic interface for implementation of real-time virtual fixtures," in *Proceedings of the 2nd Workshop on RGB-D: Advanced Reasoning with Depth Cameras (in conjunction with RSS 2011)*.

Saffer, Dan

2008 *Designing gestural interfaces: Touchscreens and interactive devices*, " O'Reilly Media, Inc.", ISBN: 0596554222.

Scuba, Private

2015 *Scuba Hand Signals*, Web Page, http://www.private-scuba.com/images/scuba-hand-signals.jpg.

Shneiderman, Ben

1993 "1.1 direct manipulation: a step beyond programming languages," *Sparks of innovation in human-computer interaction*, vol. 17, p. 1993.

2000 "Universal usability," *Communications of the ACM*, vol. 43, 5, pp. 84-91, ISSN: 0001-0782.

Sony

2013 *PlayStation Eye Information*, Web Page, https://support.us.playstation.com/app/answers/detail/a_id/372/.

Spielberg, Steven, Philip K Dick, and Janusz Kaminski

2002 *Minority Report (2002)*, Audiovisual Material.

Tomitsch, Martin, Christopher Ackad, Oliver Dawson, Luke Hespanhol, and Judy Kay

2014 "Who cares about the content? An analysis of playful behaviour at a public display," in *Proceedings of The International Symposium on Pervasive Displays*, ACM, p. 160, ISBN: 1450329527.

Tractinsky, Noam

1997 "Aesthetics and apparent usability: empirically assessing cultural and methodological issues," in *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, ACM, pp. 115-122.

Unity

2015 "Unity," https://unity3d.com/.

Usability.gov

2015 , Web Page, http://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html.

Uscrow

2015 *Tactical Hand Signals*, Web Page, http://uscrow.org/wp-content/uploads/2013/04/Tactical-Hand-Signals-968x1024.jpg.

Wobbrock, Jacob O, Htet Htet Aung, Brandon Rothrock, and Brad A Myers

2005 "Maximizing the guessability of symbolic input," in *CHI'05 extended abstracts on Human Factors in Computing Systems*, ACM, pp. 1869-1872, ISBN: 1595930027.

Wobbrock, Jacob O, Meredith Ringel Morris, and Andrew D Wilson

2009 "User-defined gestures for surface computing," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, pp. 1083-1092, ISBN: 1605582468.

Yoshioka, Keiko

2005 *Linguistic and gestural introduction and tracking of referents in L1 and L2 discourse*, Web Page, https://www.rug.nl/research/portal/files/2931153/Yoshioka-3.Ch2.pdf.

A APPENDIX

A.1 GUESSABILITY STUDY RUNSHEET

User Experiment Runsheet
(with dialogue)

Setup:

- Start the video/audio recording by the room's conference camera.
- Get ready the Kinect/webcam system to record each task.
- Get participant's consent form signed.
- Record the randomized task order for each participant on the questionnaire sheet.

Intro:

- **"This user study aims to explore user interaction with a hand gesture interface. These systems can be used in a variety of ways, for example as public information displays. For this study there are 10 tasks to complete. For each tasks, an animation will play on the screen. You will first discuss what gestures or movement you think will be appropriate for each task. I will then replay the animation three times and you will act out the movement. After each task you will complete a short questionnaire. There is a practice level at the start to get you familiarized with the process."**

For each task:

- Read the dialogue introduction.
- Play the task animation.
- Get the participant to discuss their gesture/movement.
- Restart the level (Press delete/backspace)
- Start the Kinect recording
- Play the task animation for a total of 3 times, encouraging the participant to act out the gesture/animation.
- Stop the Kinect recording
- Ask the participant to fill in the task questionnaire.

- Level P- Practice Task:

    - **"The aim of this is for you to practice interacting with the system. There is a wheel, similar to a lottery wheel which you can spin clockwise and anti-clockwise."**

    - Pressing the spacebar once will start the wheel spinning clockwise 270 degrees.
    - Pressing the spacebar again will spin the wheel anti-clockwise 180 degrees
    - The animations will replay on further presses.

- Level 1- Pointing Task:

    - **"For this level the aim is to point to the alien. The red target icon is simulating your pointing area. The alien will move first, then your gesture will simulate the movement of the target icon. The alien will move to 5 different position on the screen. The target icon will follow after a second lag, so take your time between the positions."**

    - Pressing the spacebar once will start the animation.
    - The object will move to a position on the screen, followed by the cursor icon.
    - Object will automatically move to 5 various position in total.

- Level 2- Selection Task:

    - **"For this level you have to select the yellow box. Make a gesture to select the box. The box will animate briefly to confirm your selection. There will be 5 different box to select.**"

    - Pressing the spacebar once will start the animation.
    - The cursor will move to the target, animate into a selection state, the target will disappear, and a new target will appear.
    - There is a total of 5 cycles (i.e. 5 targets).

- Level 3- Dragging Task:

    - **"For this level you will need to drag the alien to the flag icon. You will first need to gesture make the alien into a draggable state, drag it to the flag, then release the alien into its normal state."**

    - Pressing the spacebar once will start the animation.
    - The object will appear selected. Then the object will move to the target, then appear de-selected.
    - There is a total of 3 cycles.

- Level 4- Dichotomous Selection:

    - "**For this level you have to choose two option. Make a gesture to signal "YES" and another to signal "NO"**

- Level 5- Horizontal Scrolling:

    - "**For this level you will have to scroll horizontally. The object will move one direction 3 times, then move in the other direction 3 times. You will have to make a gesture that simulates the movement."**

- Pressing the spacebar once will start the animation.
- Objects will be translated to the right 3 times, then translated to the left 3 times.

- Level 6- Vertical Scrolling:

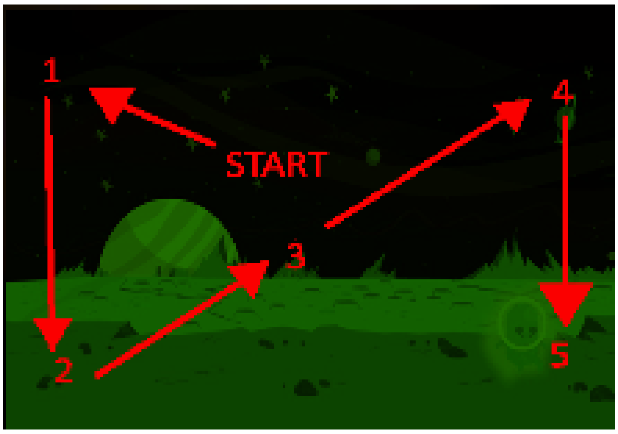
  - “**For this level you will have to scroll vertically. The object will move one direction 3 times, then move in the other direction 3 times. You will have to make a gesture that simulates the movement.”**

  - Pressing the spacebar once will play the animation.
  - Objects will be translated up 3 times, then translated down 3 times.

- Level 7- 2D Panning:

  - “**For this level you will have to explore a picture. A diagram of the picture is shown, and there are 5 areas you need to move to on the picture, in the order as shown. Simulate a gesture that would do this.”**

  - Pressing the spacebar once will play the animation.
  - The picture area will be revealed in 5 steps, as shown in the diagram below:



- Level 8- Zooming:

  - “**For this level you will have to zoom in and out on the picture. The picture will zoom in 3 times, then zoom out 3 times. You will have to make a gesture that simulates the movement.”**

  - Pressing the spacebar once will start the animation.
  - The picture will zoom in 3 times, then zoom out 3 times.

- Level 9- 3D Rotation:

    - “**For this level you will have to rotate the cube in 4 different directions- upwards, downwards, anticlockwise then clockwise. The cube will pause in between each rotation. You will have to make a gesture that simulates the movement.”**

    - Pressing the spacebar once will start the animation.
    - The cube will move in the following directions for 3 seconds with a 1 second rest in between each movement:
        1- cube rotates upwards
        2- cube rotates downwards
        3- cube rotates anticlockwise
        4- cube rotates clockwise

- Level 0- Bonus Selection Task:

    - **“For this level you have to select the alien. The target is simulating your pointing area. You have to move it to the alien, and make a gesture to select the alien. The cursor will animate briefly to confirm your selection. There will be 5 alien in various positions.**”
    - Pressing the spacebar once will start the animation.
    - The cursor will move to the target, then animate into a selection state.
    - The target will disappear, and a new target will appear.
    - There is a total of 5 cycles (i.e. 5 targets).

Once all tasks are completed:

- Get the participant to complete the post experimental questionaire
- Quickly skim through the questionaire to ensure comments are legible and prompt for clarification if necessary.
- Give the participants their cafe voucher and fill in the record sheet.

Misc.

Setting up computers:

- Conference camera:
    - This is linked to the Mac (user= multimedia)
    - Click the user study icon on desktop
    - Click the user button on keyboard
    - Press record
    - When finished, Press stop.
    - Open recorded file in Turbo program. Edit and compress video.

- Kinect and webcam recording:
    - Start the program by clicking the icon
    - Press record when tasks are being acted out
    - When finished, press stop.
    - Move the webcam data form the spare drive to the main drive.
    - Move participant files into one folder. Label the tasks and participant number.

A.2 GUESSABILITY STUDY CONSENT FORM

Human Interface Technology Laboratory New Zealand

University of Canterbury Private Bag 4800 Christchurch Telephone (64 3) 364 2349 Fax (64 3) 364 2095 Website www.hitlabnz.org Email info@hitlabnz.org

HITLabNZ

INFORMATION SHEET

RESEARCH STUDY: User-Defined Gestures for Interacting with Large Screen Displays

RESEARCHERS: Jonathan Wong, Hyungon Kim, Gun Lee, and Mark Billinghurst.

INTRODUCTION

You are invited to take part in a game design research study. Before you decide to be part of this study, you need to understand the risks and benefits. This consent form provides information about the research study. A staff member will be available to answer your questions and provide further explanations. If you agree to take part in the research study, you will be asked to sign to this consent form.

PURPOSE

The purpose of this study is to identify user friendly and intuitive gestures for interacting with large screen displays.

PROCEDURE

The study will follow the procedure outlined as below:

1. The participant reads and signs the informed consent form.

2. The participant answers to a questionnaire on demographic information and his/her previous experience with using computer interfaces.

3. The researcher explains the study setup and experimental tasks for the participant to perform during the study.

4. The participant performs the experimental tasks including:

   - Watching an animation on the screen that describes the interaction to define a gesture for.
   - Verbally describing his/her idea on the appropriate gesture for the given interaction.
   - Acting out the gesture to perform the given interaction (repeated for 3 times).
   - Rating the usability of the gesture by answering to a questionnaire.

   * While performing the gestures, the participant’s motion and verbal description of the gesture will be recorded.
   * The participant will repeat the tasks above for the provided set of interaction.

5. The participant answers to a questionnaire asking for feedback on the overall study.

6. The participant gets interviewed by the researcher on overall experience with gesture

The whole procedure will take approximately 50 minutes.

RISKS/DISCOMFORTS

Risks are minimal in this study. As you will be asked to act out gestures defined by yourself, it is expected that the experiment will involve physical movement of your body which could cause you feel tired or uncomfortable. However, as the level of physical activity will be within the range of everyday life activities, we do not expect any injury to come upon any of the participants.

CONFIDENTIALITY

All data obtained from participants will be kept confidential. In publications (e.g. Thesis, a public document which will be available through the UC Library), we will mainly report the results in an aggregate format: reporting only combined results and never reporting individual ones. In case of reporting quotes of the participants from the interviews, we will keep the source anonymous. All recordings will be concealed, and no one other than the researchers will have access to them. The data will be kept securely for a minimum period of 5 years and will be destroyed after completion of the research project.

PARTICIPATION

Participation in this research study is completely voluntary. You have the right to withdraw at anytime or refuse to participate entirely.

COMPENSATION

Upon completion of participation in the study, the participant will receive a $5 gift voucher.

APPROVAL OF THIS STUDY

This study has been reviewed and approved by the Human Interface Technology (HIT Lab NZ) and the University of Canterbury Human Ethics Committee Low Risk Approval process.

QUESTIONS

If you have questions regarding this study, please contact the researchers at the HIT Lab NZ:
Jonathan Wong (jonathan.wong@pg.canterbury.ac.nz)
Hyungon Kim (hyungon.kim@pg.canterbury.ac.nz)
Dr. Gun Lee (gun.lee@canterbury.ac.nz)
Prof. Mark Billinghurst (mark.billinghurst@canterbury.ac.nz)

Please take this information sheet with you when you leave.

PARTICIPANT CONSENT FORM

RESEARCH STUDY: User-Defined Gestures for Interacting with Large Screen Displays

RESEARCHERS: Jonathan Wong (jonathan.wong@pg.canterbury.ac.nz), Hyungon Kim (hyungon.kim@pg.canterbury.ac.nz)

SUPERVISORS: Prof. Mark Billinghurst (mark.billinghurst@canterbury.ac.nz), Dr. Gun Lee (gun.lee@canterbury.ac.nz)

I have been given a full explanation of this project and have had the opportunity to ask questions. I understand what is required of me if I agree to take part in the research.

I understand that participation is voluntary and I may withdraw at any time without penalty. Withdrawal of participation will also include the withdrawal of any information I have provided should this remain practically achievable.

I understand that any information or opinions I provide will be kept confidential to the researcher and the administrators of the research project and that any published or reported results will not identify the participants. I understand that a thesis is a public document and will be available through the UC Library

I understand that all data collected for the study will be kept in locked and secure facilities and/or in password protected electronic form and will be destroyed after five years.

I understand the risks associated with taking part and how they will be managed.

I understand that I am able to receive a report on the findings of the study by contacting the researcher at the conclusion of the project.

I understand that I can contact the researchers or supervisors listed above for further information. If I have any complaints, I can contact the Chair of the University of Canterbury Human Ethics Committee, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz)

By signing below, I agree to participate in this research project, and I authorize recordings or other materials taken from this study used for scientific purposes, and I consent to publication of the results of the study.

______________________	______________	________
Participant (Print name)	Signature	Date

A.3 GUESSABILITY STUDY QUESTIONNAIRE

Task Order: ___ - ___ - ___ - ___ - ___ - ___ - ___ - ___ - ___ - ___

Pre-experimental Questionnaire

Age: ____________ ☐ Male / ☐ Female

Please check on ONE answer, unless it is described otherwise.

1. What is your first language?

____

2. If applicable, what is your second language?

____

I started using my second language since I was _____ years old.

3. Have you used free hand gesture based interface before?

☐ Not at all

☐ Couple of times a year

☐ Couple of times in a month

☐ Couple of times in a week

☐ Every day

4. Have you played XBOX Kinect motion games before?

☐ Not at all

☐ Couple of times a year

☐ Couple of times in a month

☐ Couple of times in a week

☐ Every day

5. Have you played Nintendo Wii or Sony MOVE motion games before?

☐ Not at all

☐ Couple of times a year

☐ Couple of times in a month

☐ Couple of times in a week

☐ Every day

6. If you have played motion games other than those mentioned above, what are they? ______________________________
And how often did you play?

☐ Not at all

☐ Couple of times a year

☐ Couple of times in a month

☐ Couple of times in a week

☐ Every day

7. Have you used an Augmented Reality (AR) app/interface before?

☐ I am not aware of what AR is.

☐ Not at all

☐ Couple of times a year

☐ Couple of times in a month

☐ Couple of times in a week

☐ Every day

8. Which hand do you usually use for pointing or making gestures?

☐ Left

☐ Right

☐ Both

Please check on the box based on how much you agree with each statement.

9. I consider myself using gestures a lot in everyday life.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

Thank you! Please wait for further instruction.

Per-Task Questionnaire Task# _____

Please check on the box based on how much you agree with each statement.

1. I was able to **perform the gesture well**.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

2. The gesture is **easy to perform.**

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

3. The gesture will be **easy to learn** by others.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

4. The gesture is **intuitive**.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

5. The gesture feels **natural**.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

6. The gesture is **efficient** to do the given task.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

7. The gesture is **mentally stressful** to perform.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

8. The gesture is **physically stressful** to perform.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

9. The gesture can be **easily recognized** by another person.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

10. The gesture can be **confusing** with another gesture.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

Post-experimental Questionnaire

Please check on the box based on how much you agree with each statement.

1. In overall, the whole set of gestures is **well defined**.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

2. The gestures as a whole set are **easy to perform**.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

3. The gestures as a whole set will be **easy to learn** by others.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

4. The gestures as a whole set is **intuitive**.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

5. The gestures as a whole set feels **natural**.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

6. Some of the gestures can be **confusing** with another gesture.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

7. If given more time, I would modify the gestures to be more useful.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

8. If you would modify or improve some of the gestures, which gesture(s)/task(s) do you want to modify, and how?

Assuming that a public information display can recognize user's gestures, please answer to the following questions.

9. I would feel comfortable to interact with the information display **using gestures in public space**.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

10. I would be **more encouraged to use** the information display if using gestures to interact with it.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

11. What kind of **places** do you think such gesture interactive public information display would be useful?

12. What kind of **potential problems** could there be when interacting with public information displays using gestures?

13. What other types of interaction (controlling) methods do you think would be useful for interacting with public information displays?

Thank you for your participation!

A.4 EXPERIMENT CONSENT FORM

Human Interface Technology Laboratory New Zealand

University of Canterbury Private Bag 4800 Christchurch Telephone (64 3) 364 2349 Fax (64 3) 364 2095 Website www.hitlabnz.org Email info@hitlabnz.org

HITLabNZ

INFORMATION SHEET

RESEARCH STUDY: Gesture based Interaction for Augmented Virtual Mirrors.

RESEARCHERS: Jonathan Wong, Christoph Bartneck, Gun Lee, and Mark Billinghurst.

INTRODUCTION

You are invited to take part in a gesture interface research study. Before you decide to be part of this study, you need to understand the risks and benefits. This consent form provides information about the research study. A staff member will be available to answer your questions and provide further explanations. If you agree to take part in the research study, you will be asked to sign to this consent form.

PURPOSE

The purpose of this study is to identify user friendly and intuitive gestures for interacting with large screen displays.

PROCEDURE

The study will follow the procedure outlined as below:

1. The participant reads and signs the informed consent form.

2. The participant answers to a questionnaire on demographic information and his/her previous experience with using computer interfaces.

3. The researcher explains the study setup and experimental tasks for the participant to perform during the study.

4. The participant performs the experimental tasks which may include:

   - Watching an animation on the screen that describes the interaction to define a gesture for.
   - Performing a gesture to perform selected interactions on the screen.
   - Rating the usability of the gesture by answering to a questionnaire.

5. The participant answer a questionnaire asking for feedback on the overall study.

The whole procedure will take approximately 40 minutes.

University of Canterbury Private Bag 4800 Christchurch Telephone (64 3) 364 2349 Fax (64 3) 364 2095 Website www.hitlabnz.org Email info@hitlabnz.org

RISKS/DISCOMFORTS

Risks are minimal in this study. As you will be asked to act out gestures defined by yourself, it is expected that the experiment will involve physical movement of your body which could cause you feel tired or uncomfortable. However, as the level of physical activity will be within the range of everyday life activities, we do not expect any injury to come upon any of the participants.

CONFIDENTIALITY

All data obtained from participants will be kept confidential. In publications (e.g. Thesis, a public document which will be available through the UC Library), we will mainly report the results in an aggregate format: reporting only combined results and never reporting individual ones. In case of reporting quotes of the participants from the interviews, we will keep the source anonymous. All recordings will be concealed, and no one other than the researchers will have access to them. The data will be kept securely for a minimum period of 5 years and will be destroyed after completion of the research project.

PARTICIPATION

Participation in this research study is completely voluntary. You have the right to withdraw at anytime or refuse to participate entirely.

COMPENSATION

Upon completion of participation in the study, the participant will receive a $5 gift voucher.

APPROVAL OF THIS STUDY

This study has been reviewed and approved by the Human Interface Technology (HIT Lab NZ) and the University of Canterbury Human Ethics Committee Low Risk Approval process.

QUESTIONS

If you have questions regarding this study, please contact the researchers at the HIT Lab NZ:

Jonathan Wong (jonathan.wong@pg.canterbury.ac.nz)
Christoph Bartneck (christoph.bartneck@canterbury.ac.nz)
Gun Lee (gun.lee@canterbury.ac.nz)
Mark Billinghurst (mark.billinghurst@canterbury.ac.nz)

Please take this information sheet with you when you leave.

Human Interface Technology Laboratory New Zealand

University of Canterbury Private Bag 4800 Christchurch Telephone (64 3) 364 2349 Fax (64 3) 364 2095 Website www.hitlabnz.org Email info@hitlabnz.org

HITLabNZ

PARTICIPANT CONSENT FORM

RESEARCH STUDY: Gesture based Interaction for Augmented Virtual Mirrors.

RESEARCHER: Jonathan Wong

SUPERVISORS: Dr. Christoph Bartneck (christoph.bartneck@canterbury.ac.nz), Dr. Gun Lee (gun.lee@canterbury.ac.nz), Prof. Mark Billinghurst (mark.billinghurst@canterbury.ac.nz),

I have been given a full explanation of this project and have had the opportunity to ask questions. I understand what is required of me if I agree to take part in the research.

I understand that participation is voluntary and I may withdraw at any time without penalty. Withdrawal of participation will also include the withdrawal of any information I have provided should this remain practically achievable.

I understand that any information or opinions I provide will be kept confidential to the researcher and the administrators of the research project and that any published or reported results will not identify the participants. I understand that a thesis is a public document and will be available through the UC Library

I understand that all data collected for the study will be kept in locked and secure facilities and/or in password protected electronic form and will be destroyed after five years.

I understand the risks associated with taking part and how they will be managed.

I understand that I am able to receive a report on the findings of the study by contacting the researcher at the conclusion of the project.

I understand that I can contact the researchers or supervisors listed above for further information. If I have any complaints, I can contact the Chair of the University of Canterbury Human Ethics Committee, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz)

By signing below, I agree to participate in this research project, and I authorize recordings or other materials taken from this study used for scientific purposes, and I consent to publication of the results of the study.

____________________________ __________________ ________

Participant (Print name) Signature Date

A.5 EXPERIMENT QUESTIONNAIRE

Pre-experimental Questionnaire

Age: _____________ ☐ Male / ☐ Female

Please check on ONE answer, unless it is described otherwise.

1. What is your first language?

____

2. If applicable, what is your second language?

____

I started using my second language since I was ____ years old.

3. Have you used free hand gesture based interface before?

☐ Not at all

☐ Couple of times a year

☐ Couple of times in a month

☐ Couple of times in a week

☐ Every day

4. Have you played XBOX Kinect motion games before?

☐ Not at all

☐ Couple of times a year

☐ Couple of times in a month

☐ Couple of times in a week

☐ Every day

5. Have you played Nintendo Wii or Sony MOVE motion games before?

☐ Not at all

☐ Couple of times a year

☐ Couple of times in a month

☐ Couple of times in a week

☐ Every day

6. If you have played motion games other than those mentioned above, what are they? ______________________________
and how often did you play?

☐ Not at all

☐ Couple of times a year

☐ Couple of times in a month

☐ Couple of times in a week

☐ Every day

7. Have you used an Augmented Reality (AR) app/interface before?

☐ I am not aware of what AR is.

☐ Not at all

☐ Couple of times a year

☐ Couple of times in a month

☐ Couple of times in a week

☐ Every day

8. Which hand do you usually use for pointing or making gestures?

☐ Left

☐ Right

☐ Both

Please check on the box based on how much you agree with each statement.

9. I consider myself using gestures a lot in everyday life.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

Thank you! Please wait for further instruction.

Per-Task Questionnaire Task# _____

Please check on the box based on how much you agree with each statement.

1. I was able to **perform the gesture well**.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

2. The gesture is **easy to perform.**

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

3. The gesture will be **easy to learn** by others.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

4. The gesture is **intuitive**.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

5. The gesture feels **natural**.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

6. The gesture is **fun** to perform.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

7. The gesture feels **accurate**.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

8. The gesture is **efficient** to do the given task.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

9. The gesture is **mentally stressful** to perform.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

10. The gesture is **physically stressful** to perform.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

11. The gesture can be **easily recognized** by another person.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

12. The gesture can be **easily recognized** by the computer.

Not at all 1	2	3	Moderately 4	5	6	Extremely 7
☐	☐	☐	☐	☐	☐	☐

Post-experimental Questionnaire

Rank the conditions with the following letters:

H- Hover

E- Extension

P-Pointing

1. Which condition was the **easiest to perform**?

Easiest		←→		Hardest
☐		☐		☐

2. Which condition was the **most intuitive**?

Least intuitive		←→		Most Intuitive
☐		☐		☐

3. Which condition was the most **accurate?**

Least accurate		←→		Most accurate
☐		☐		☐

4. Which condition was the most **fun** to use?

Least fun		←→		Most fun
☐		☐		☐

5. Which condition did you **prefer overall**?

Least preferred		←→		Most preferred
☐		☐		☐

6. Comment on the different conditions (**H**overing, **E**xtension and **P**ointing). E.g. what was good about the conditions? What did you dislike? What would make it better?

Thank you for your participation!